Almost-sure hedging with permanent price impact

B. Bouchard* and G. Loeper† and Y. Zou ‡

November 3, 2015

Abstract

We consider a financial model with permanent price impact. Continuous time trading dynamics are derived as the limit of discrete rebalancing policies. We then study the problem of super-hedging a European option. Our main result is the derivation of a quasi-linear pricing equation. It holds in the sense of viscosity solutions. When it admits a smooth solution, it provides a perfect hedging strategy.

Keywords: Hedging, Price impact.

AMS 2010 Subject Classification: 91G20; 93E20; 49L20

Introduction

Two of the fundamental assumptions in the Black and Scholes approach for option hedging are that the price dynamics are unaffected by the hedger's behaviour, and that he can trade unrestricted amounts of asset at the instantaneous value of the price process. In other words, it relies on the absence of market impact and of liquidity costs or liquidity constraints. This work addresses the problem of option hedging under a price dynamics model that incorporates directly the hedger's trading activity, and hence that violates those two assumptions.

*CEREMADE, Université Paris Dauphine and CREST-ENSAE. Research supported by ANR Liquirisk and Investissements d'Avenir (ANR-11-IDEX-0003/Labex Ecodec/ANR-11-LABX-0047).

†BNP-Paribas and FiQuant - Chaire de finance quantitative

‡CEREMADE, Université Paris Dauphine and CREST-ENSAE.

In the literature, one finds numerous studies related to this topic. Some of them incorporate liquidity costs but no price impact, the price curve is not affected by the trading strategy. In the setting of [6], this does not affect the super-hedging price because trading can essentially be done in a bounded variation manner at the marginal spot price at the origine of the curve. However, if additional restrictions are imposed on admissible strategies, this leads to a modified pricing equation, which exhibits a quadratic term in the second order derivative of the solution, and renders the pricing equation fully non-linear, and even not unconditionally parabolic, see [7] and [20]. Another branch of literature focuses on the derivation of the price dynamics through clearing condition. In the papers [9], [16], [15], the authors work on supply and demand curves that arise from "reference" and "program" traders (i.e. option hedgers) to establish a modified price dynamics, but do not take into account the liquidity costs, see also [12]. This approach also leads to non-linear pde's, but the non-linearity comes from a modified volatility process rather than from a liquidity cost source term. Finally, the series of papers [17], [19], [14] address the liquidity issue indirectly by imposing bounds on the "gamma" of admissible trading strategies, no liquidity cost or price impact are modeled explicitly.

More recently, [13] and [1] have considered a novel approach in which the price dynamic is driven by the sum of a classical Wiener process and a (locally) linear market impact term. The linear market impact mechanism induces a modified volatility process, as well as a non trivial average execution price. However, the trader starts his hedging with the correct position in stocks and does not have to unwind his final position (this corresponds to "covered" options with delivery). Those combined effects lead to a fully non-linear pde giving the exact replication strategy, which is not always parabolic depending on the ratio between the instantaneous market impact (liquidity costs) and permanent market impact.

In this paper we build on the same framework as [13], in the case where the instantaneous market impact equals the permanent impact (no relaxation effect), and go one step further by considering the effect of (possibly) unwinding the portfolio at maturity, and of building the initial portfolio. Consequently the spot "jumps" at initial time when building the hedge portfolio, and at maturity when unwinding it (depending on the nature of the payoff - delivery can also be made in stocks). In this framework, we find that the optimal super-replication strategy follows a modified quasi-linear Black and Scholes pde. Although the underlying model is similar to the one proposed by the second author [13], the pricing pde is therefore fundamentally different (quasi-linear vs fully non-linear).

Concerning the mathematical approach, while in [13] the author focused on exhibiting an exact replication strategy by a verification approach, in this work we follow a stochastic target approach and derive the pde from a dynamic programming principle. The difficulty is that, because of the market impact mechanism, the state process must be described by the asset price and the hedger's portfolio (i.e. the amount of risky asset detained by the hedger) and this leads to a highly singular control problem. It is overcome by a suitable change of variable which allows one to reduce to a zero initial position in the risky asset and state a version of the geometric dynamic programming principle in terms of the post-portfolio liquidation asset price process: the price that would be obtained if the trader was liquidating his position immediately.

The paper is organized as follows. In Section 1, we present the impact rule and derive continuous time trading dynamics as limits of discrete time rebalancing policies. The super-hedging problem is set in Section 2 as a stochastic target problem. We first prove a suitable version of the geometric dynamic programming and then derive the corresponding pde in the viscosity solution sense. Uniqueness and regularity are established under suitable assumptions. We finally further discuss the case of a constant impact coefficients, to provide a better understanding of the "hedging strategy".

General notations. Given a function ϕ, we denote by ϕ' and ϕ'' its first and second order derivatives if they exist. When ϕ depends on several arguments, we use the notations $\partial_x \phi$, $\partial^2_{xx}\phi$ to denote the first and second order partial derivatives with respect to its x-argument, and write $\partial^2_{xy}\phi$ for the cross second order derivative in its (x, y)-argument.

All over this paper, Ω is the canonical space of continuous functions on $\mathbb{R}_+$ starting at 0, $\mathbb{P}$ is the Wiener measure, W is the canonical process, and $\mathbb{F} = (\mathcal{F}_t)_{t\geq 0}$ is its augmented raw filtration. All random variables are defined on $(\Omega, \mathcal{F}_\infty, \mathbb{P})$. $\mathbf{L}_0$ (resp. $\mathbf{L}_2$) denotes the space of (resp. square integrable) $\mathbb{R}^n$-valued random variables, while $\mathbf{L}_0^\lambda$ (resp. $\mathbf{L}_2^\lambda$) stands for the collection of predictable $\mathbb{R}^n$-valued processes ϑ (resp. such that $\|\vartheta\|_{\mathbf{L}_2^\lambda} := \mathbb{E}[\int_0^\infty |\vartheta_s|^2 ds]^{\frac{1}{2}}$). The integer $n \geq 1$ is given by the context and $|x|$ denote the Euclidean norm of $x \in \mathbb{R}^n$.

1 Portfolio and price dynamics

This section is devoted to the derivation of our model with continuous time trading. We first consider the situation where a trading signal is given by a continuous Itô process and the position in stock is rebalanced in discrete time. In this case, the dynamics of the stock price and the wealth process are given according to our impact rule. A first continuous time trading dynamic is obtained by letting the time between two consecutive trades vanish. Then, we incorporate jumps as the limit of continuous trading on a short time horizon.

We restrict here to a single stock market. This is only for simplicity, the extension to a multi-dimensional market is just a matter of notations.

1.1 Impact rules

We model the impact of a strategy on the price process through an impact function f: the price variation du to buying a (infinitesimal) number $\delta \in \mathbb{R}$ of shares is $\delta f(x)$, if the price of the asset is x before the trade. The cost of buying the additional δ units is given by

$$\delta x + \frac{1}{2}\delta^2 f(x) = \delta \int_0^\delta \frac{1}{\delta}(x + f(x)\iota)d\iota,$$

in which

$$\int_0^\delta \frac{1}{\delta}(x + f(x)\iota)d\iota$$

should be interpreted as the average cost for each additional unit. Between two times of trading $\tau_1 \leq \tau_2$, the dynamics of the stock is given by the strong solution of the stochastic differential equation

$$dX_t = \mu(X_t)dt + \sigma(X_t)dW_t. \tag{1.1}$$

All over this paper, we assume that

$$\begin{array}{c} f \in C_b^2 \text{ and is (strictly) positive,} \\ (\mu, \sigma, \sigma^{-1}) \text{ is Lipschitz and bounded.} \end{array} \tag{H1}$$

Remark 1.1. *a. We restrict here to an impact rule which is linear in the size of the order. However, note that in the following it will only be applied to order of infinitesimal size (at the limit). One would therefore obtain the same final dynamics* (1.23)-(1.24) *below by considering a more general impact rule* $\delta \mapsto F(x, \delta)$ *whenever is satisfies* $F(x, 0)= \partial^2_{\delta\delta}F(x, 0) = 0$ *and* $\partial_\delta F(x, 0) = f(x)$*. See Remark 1.2 below. Otherwise stated, for our analysis, we only need to consider the value and the slope at* $\delta = 0$ *of the impact function.*

b. A typical example of such a function is $F = \Delta\mathrm{x}$ *where*

$$\Delta\mathrm{x}(x, \delta) := \mathrm{x}(x, \delta) - x\,, \tag{1.2}$$

with $\mathrm{x}(x, \cdot)$ *defined as the solution of*

$$\mathrm{x}(x, \cdot) \;=\; x + \int_0^\cdot f(\mathrm{x}(x, s))ds. \tag{1.3}$$

The curve x *has a natural interpretation. For an order of small size* $\Delta\iota$*, the stock price jumps from* x *to* $x + \Delta\iota f(x) \simeq \mathrm{x}(x, \Delta\iota)$*. Passing another order of size* $\Delta\iota$

makes it move again to approximately $\mathrm{x}(\mathrm{x}(x,\Delta\iota),\Delta\iota)=\mathrm{x}(x,2\Delta\iota)$*, etc. Passing to the limit* $\Delta\iota\to 0$ *but keeping the total trade size equal to* δ *provides asymptotically a price move equal to* $\Delta\mathrm{x}(x,\delta)$.

This specific curve will play a central role in our analysis, see Section 1.3.

1.2 Discrete rebalancing from a continuous signal and continuous time trading limit

We first consider the situation in which the number of shares the trader would like to hold is given by a continuous Itô process Y of the form

$$Y = Y_0 + \int_0^{\cdot} b_s ds + \int_0^{\cdot} a_s dW_s, \tag{1.4}$$

where

$$(a,b) \in \mathcal{A} := \cup_k \mathcal{A}_k,$$
$$\mathcal{A}_k := \{(a,b) \in \mathbf{L}_0^{\lambda} : |(a,b)| \le k \ dt \times d\mathbb{P} - \text{a.e.}\} \text{ for } k > 0.$$

In order to derive our continuous time trading dynamics, we consider the corresponding discrete time rebalancing policy set on a time grid

$$t_i^n := iT/n,\ i=0,\dots,n,\ n\ge 1,$$

and then pass to the limit $n\to\infty$.

If the trader only changes the composition of his portfolio at the discrete times t_i^n, then he holds $Y_{t_i^n}$ stocks on each time interval $[t_i^n, t_{i+1}^n)$. Otherwise stated, the number of shares actually held at $t\le T$ is

$$Y_t^n := \sum_{i=0}^{n-1} Y_{t_i^n} \mathbf{1}_{\{t_i^n \le t < t_{i+1}^n\}} + Y_T \mathbf{1}_{\{t=T\}} \tag{1.5}$$

and the number of purchased shares is

$$\delta_t^n := \sum_{i=1}^{n} \mathbf{1}_{\{t=t_i^n\}} (Y_{t_i^n} - Y_{t_{i-1}^n}).$$

Given our impact rule, the corresponding dynamics for the stock price process is

$$X^n = X_0 + \int_0^{\cdot} \mu(X_s^n) ds + \int_0^{\cdot} \sigma(X_s^n) dW_s + \sum_{i=1}^{n} \mathbf{1}_{[t_i^n, T]} \delta_{t_i^n}^n f(X_{t_i^n -}^n), \tag{1.6}$$

in which X_0 is a constant.

To describe the portfolio process, we provide the dynamics of the sum V^n of the amount of cash held and the potential amount $Y^n X^n$ associated to the position in stocks:

$$V^n = \text{cash position } + Y^n X^n. \tag{1.7}$$

Observe that this is not the liquidation value of the portfolio, except when $Y^n = 0$, as the liquidation of Y^n stocks will have an impact on the market and does not generate a gain equal to $Y^n X^n$. However, if we keep Y^n in mind, the couple (V^n, Y^n) gives the exact composition in cash and stocks of the portfolio. By a slight abuse of language, we call V^n the portfolio value or wealth process.

Assuming that the risk free rate is zero (for ease of notations), its dynamics is given by

$$V^n = V_0 + \int_0^\cdot Y^n_{s-} dX^n_s + \sum_{i=1}^n \mathbf{1}_{[t^n_i, T]} \frac{1}{2} (\delta^n_{t^n_i})^2 f(X^n_{t^n_i -}), \tag{1.8}$$

or equivalently

$$\begin{aligned} V^n = V_0 &+ \sum_{i=1}^n \mathbf{1}_{[t^n_{i-1}, T]} Y_{t^n_{i-1}} (X^n_{\cdot \wedge t^n_i -} - X^n_{t^n_{i-1}}) \\ &+ \sum_{i=1}^n \mathbf{1}_{[t^n_i, T]} \left[\frac{1}{2} (\delta^n_{t^n_i})^2 f(X^n_{t^n_i -}) + Y_{t^n_{i-1}} \delta^n_{t^n_i} f(X^n_{t^n_i -}) \right], \end{aligned} \tag{1.9}$$

in which $V_0 \in \mathbb{R}$. Let us comment this formula. The first term on the right-hand side corresponds to the evolution of the portfolio value strictly between two trades ; it is given by the number of shares held multiplied by the price increment. When a trade of size $\delta^n_{t^n_i}$ occurs at time t^n_i, the cost of buying the stocks is $2^{-1} (\delta^n_{t^n_i})^2 f(X^n_{t^n_i -}) + \delta^n_{t^n_i} X^n_{t^n_i -}$ but it provides $\delta^n_{t^n_i}$ more stocks, on top of the $Y^n_{t^n_i -} = Y_{t^n_{i-1}}$ units that are already in the portfolio. After the price's move generated by the trade, the stocks are evaluated at $X^n_{t^n_i}$. The increment in value du to the price's move and the additional position is therefore $\delta^n_{t^n_i} X^n_{t^n_i} + Y^n_{t^n_i -} (X^n_{t^n_i} - X^n_{t^n_i -})$. Since $X^n_{t^n_i} - X^n_{t^n_i -} = \delta^n_{t^n_i} f(X^n_{t^n_i -})$, we obtain (1.9), a compact version of which is given in (1.8).

Our continuous time trading dynamics are obtained by passing to the limit $n \to \infty$, i.e. by considering faster and faster rebalancing strategies.

Proposition 1.1. *Let $Z := (X, Y, V)$ where Y is defined as in* (1.4) *for some $(a, b) \in \mathcal{A}$, and (X, V) solves*

$$X = X_0 + \int_0^\cdot \sigma(X_s) dW_s + \int_0^\cdot f(X_s) dY_s + \int_0^\cdot (\mu(X_s) + a_s (\sigma f')(X_s)) ds \tag{1.10}$$

and

$$V = V_0 + \int_0^\cdot Y_s dX_s + \frac{1}{2}\int_0^\cdot a_s^2 f(X_s) ds. \tag{1.11}$$

Let $Z^n := (X^n, Y^n, V^n)$ be defined as in (1.6)-(1.5)-(1.8). Then, there exists a constant $C > 0$ such that

$$\sup_{[0,T]} \mathbb{E}\left[|Z^n - Z|^2\right] \le Cn^{-1}$$

for all $n \ge 1$.

Proof. This follows standard arguments and we only provide the main ideas. In all this proof, we denote by C a generic positive constant which does not depend on n nor $i \le n$, and may change from line to line. We shall use repeatedly (H1) and the fact that a and b are bounded by some constant k, in the $dt \times d\mathbb{P}$-a.e. sense.

a. The convergence of the process Y^n is obvious:

$$\sup_{[0,T]} \mathbb{E}\left[|Y^n - Y|^2\right] \le Cn^{-1}. \tag{1.12}$$

For later use, set $\Delta X^n := X - X^n$ and also observe that the estimate

$$\sup_{[t_{i-1}^n, t_i^n)} \mathbb{E}\left[|\Delta X^n|^2\right] \le \mathbb{E}\left[|\Delta X^n_{t_{i-1}^n}|^2\right](1 + Cn^{-1}) + Cn^{-1}, \tag{1.13}$$

is standard. We now set

$$\tilde{X}_t^n := X_t^n + A_t^{i,n} + B_t^{i,n}, \;\; t_{i-1}^n \le t < t_i^n,$$

where

$$\begin{aligned} A_t^{i,n} &:= \int_{t_{i-1}^n}^t f(X_s^n) dY_s + \int_{t_{i-1}^n}^t a_s (\sigma f')(X_s^n) ds \\ B_t^{i,n} &:= \int_{t_{i-1}^n}^t (Y_s - Y_{t_{i-1}^n})(\mu f' + \frac{1}{2}\sigma^2 f'')(X_s^n) ds + \int_{t_{i-1}^n}^t (Y_s - Y_{t_{i-1}^n})(\sigma f')(X_s^n) dW_s. \end{aligned}$$

Since $A_{t_i^n}^{i,n} + B_{t_i^n}^{i,n} = \delta_{t_i^n}^n f(X_{t_i^n -}^n)$, we have

$$\lim_{t \uparrow t_i^n} \tilde{X}_t^n = X_{t_i^n}^n.$$

Set $\Delta \tilde{X}^n := X - \tilde{X}^n$, $\beta^1 := bf + a\sigma f'$ and $\beta^2 := af$, so that

$$\begin{aligned} d|\Delta \tilde{X}_t^n|^2 &= 2\Delta \tilde{X}_t^n [(\mu + \beta_t^1)(X_t) - (\mu + \beta_t^1)(X_t^n)] dt \\ &\quad + [(\sigma + \beta_t^2)(X_t) - (\sigma + \beta_t^2)(X_t^n) - (Y_t - Y_{t_{i-1}^n})(\sigma f')(X_t^n)]^2 dt \\ &\quad - 2\Delta \tilde{X}_t^n (Y_t - Y_{t_{i-1}^n})(\mu f' + \frac{1}{2}\sigma^2 f'')(X_t^n) dt \\ &\quad + 2\Delta \tilde{X}_t^n [(\sigma + \beta_t^2)(X_t) - (\sigma + \beta_t^2)(X_t^n)] dW_t \\ &\quad - 2\Delta \tilde{X}_t^n (Y_t - Y_{t_{i-1}^n})(\sigma f')(X_t^n) dW_t. \end{aligned}$$

In view of (1.12)-(1.13), this implies, for $t_{i-1}^n \le t < t_i^n$,

$$\mathbb{E}\left[|\Delta \tilde{X}_t^n|^2\right] \le \mathbb{E}\left[|\Delta X_{t_{i-1}^n}^n|^2\right] + C\mathbb{E}\left[\int_{t_{i-1}^n}^t (|\Delta \tilde{X}_s^n|^2 + |X_s - X_s^n|^2 + |Y_s - Y_{t_{i-1}^n}|^2)ds\right]$$
$$\le \mathbb{E}\left[|\Delta X_{t_{i-1}^n}^n|^2\right](1 + Cn^{-1}) + C\mathbb{E}\left[\int_{t_{i-1}^n}^t |\Delta \tilde{X}_s^n|^2 ds + n^{-2}\right],$$

and therefore

$$\sup_{[t_{i-1}^n, t_i^n)} \mathbb{E}\left[|\Delta \tilde{X}^n|^2\right] \quad \le \quad \mathbb{E}\left[|\Delta X_{t_{i-1}^n}^n|^2\right](1 + Cn^{-1}) + Cn^{-2}, \tag{1.14}$$

by Gronwall's Lemma. Since $\lim_{t\uparrow t_i^n} \tilde{X}_t^n = X_{t_i^n}^n$, this shows that

$$\mathbb{E}\left[|\Delta X_{t_i^n}^n|^2\right] \le \sup_{[t_{i-1}^n, t_i^n)} \mathbb{E}\left[|\Delta \tilde{X}^n|^2\right] \le Cn^{-1} \quad \text{for all } i \le n.$$

Plugging this inequality in (1.13), we then deduce

$$\sup_{[t_{i-1}^n, t_i^n]} \mathbb{E}\left[|\Delta X^n|^2\right] \quad \le \quad Cn^{-1} \quad \text{for all } i \le n. \tag{1.15}$$

b. We now consider the difference $V - V^n$. It follows from (1.9) that

$$V_{t_i^n}^n = V_{t_{i-1}^n}^n + \int_{t_{i-1}^n}^{t_i^n} Y_{t_{i-1}^n} \mu(X_s^n) ds + \int_{t_{i-1}^n}^{t_i^n} Y_{t_{i-1}^n} \sigma(X_s^n) dW_s$$
$$+ \int_{t_{i-1}^n}^{t_i^n} \left(\frac{1}{2} a_s^2 f(X_s^n) + Y_{t_{i-1}^n} a_s (f'\sigma)(X_s^n)\right) ds + \int_{t_{i-1}^n}^{t_i^n} Y_{t_{i-1}^n} f(X_s^n) dY_s$$
$$+ \int_{t_{i-1}^n}^{t_i^n} \alpha_s^{1n} ds + \int_{t_{i-1}^n}^{t_i^n} \alpha_s^{2n} dW_s$$

where, by (1.12), α^{1n} and α^{2n} are adapted processes satisfying

$$\sup_{[t_{i-1}^n, t_i^n)} \mathbb{E}[|\alpha^{1n}|^2 + |\alpha^{2n}|^2] \le Cn^{-1}.$$

In view of (1.12)-(1.15), this leads to

$$V_{t_i^n}^n = V_{t_{i-1}^n}^n + V_{t_i^n} - V_{t_{i-1}^n} + \int_{t_{i-1}^n}^{t_i^n} \gamma_s^{1n} ds + \int_{t_{i-1}^n}^{t_i^n} \gamma_s^{2n} dW_s \tag{1.16}$$

where γ^{1n} and γ^{2n} are adapted processes satisfying

$$\sup_{[t^n_{i-1},t^n_i)} \mathbb{E}[|\gamma^{1n}|^2 + |\gamma^{2n}|^2] \leq Cn^{-1}. \tag{1.17}$$

Set

$$\tilde{V}^n_t := V^n_{t^n_{i-1}} + V_t - V_{t^n_{i-1}} + \int_{t^n_{i-1}}^t \gamma^{1n}_s ds + \int_{t^n_{i-1}}^t \gamma^{2n}_s dW_s, \ \ t^n_{i-1} \leq t < t^n_i.$$

Then, by applying Itô's Lemma to $|\tilde{V}^n_t - V_t|^2$, using (1.17) and Gronwall's Lemma, we obtain

$$\sup_{[t^n_{i-1},t^n_i)} \mathbb{E}\left[|\tilde{V}^n - V|^2\right] \leq \mathbb{E}\left[|V^n_{t^n_{i-1}} - V_{t^n_{i-1}}|^2\right](1 + Cn^{-1}) + Cn^{-2},$$

so that, by the identity $\lim_{t\uparrow t^n_i} \tilde{V}^n_t = V^n_{t^n_i}$, recall (1.16), and an induction,

$$\mathbb{E}\left[|V^n_{t^n_i} - V_{t^n_i}|^2\right] \leq Cn^{-1}, \ \ i \leq n.$$

We conclude by observing that

$$\begin{aligned}
\mathbb{E}\left[|V^n_t - V_t|^2\right] &\leq C\mathbb{E}\left[|V^n_{t^n_{i-1}} - V_{t^n_{i-1}}|^2 + |V^n_{t^n_{i-1}} - V^n_t|^2 + |V_{t^n_{i-1}} - V_t|^2\right] \\
&\leq C\left(\mathbb{E}\left[|V^n_{t^n_{i-1}} - V_{t^n_{i-1}}|^2\right] + n^{-1}\right),
\end{aligned}$$

for $t^n_{i-1} \leq t < t^n_i$. □

Remark 1.2. *If the impact function $\delta f(x)$ was replaced by a more general C^2_b one of the form $F(x,\delta)$, with $F(x,0) = \partial^2_{\delta\delta}F(x,0) = 0$, the computations made in the above proof would only lead to terms of the from $\partial_\delta F(X,0)dY$ and $a\sigma(X)\partial^2_{x\delta}F(X,0)$ in place of $f(X)dY$ and $a(\sigma f')(X)$ in the dynamics* (1.10)*. Similarly, the term $a^2 f(X)$ would be replaced by $a^2\partial_\delta F(X,0)$ in* (1.11).

1.3 Jumps and large orders splitting

We now explain how we incorporate jumps in our dynamics. Let $\mathcal{U}_k$ denote the set of random $\{0,\cdots,k\}$-valued measures ν supported by $[-k,k]\times[0,T]$ that are adapted in the sense that $t \mapsto \nu(A\times[0,t])$ is adapted for all Borel subset A of $[-k,k]$. We set

$$\mathcal{U} := \cup_{k\geq 0}\,\mathcal{U}_k.$$

Note that an element ν of $\mathcal{U}$ can be written in the form

$$\nu(A,[0,t]) = \sum_{j=1}^{k} \mathbf{1}_{\{(\delta_j,\tau_j)\in A\times[0,t]\}} \tag{1.18}$$

in which $0 \le \tau_1 < \cdots < \tau_k \le T$ are stopping times and each δ_j is a real-valued $\mathcal{F}_{\tau_j}$-random variable.

Then, given $(a,b,\nu) \in \mathcal{A} \times \mathcal{U}$, we define the trading signal as

$$Y = Y_{0-} + \int_0^{\cdot} b_s ds + \int_0^{\cdot} a_s dW_s + \int_0^{\cdot}\int \delta\nu(d\delta, ds), \tag{1.19}$$

where $Y_{0-} \in \mathbb{R}$. For later use, we let Y^c denote its continuous part, i.e. $Y^c := Y - \int_0^{\cdot}\int \delta\nu(d\delta, ds)$.

In view of the previous sections, we assume that the dynamics of the stock price and portfolio value processes are given by (1.10)-(1.11) when Y has no jump. We incorporate jumps by assuming that the trader follows the natural idea of splitting a large order δ_j into small pieces on a small time interval. This is a current practice which aims at avoiding having a too large impact, and paying a too high liquidity cost. Given the asymptotic already derived in the previous section, we can reduce to the case where this is done continuously at a constant rate δ_j/ε on $[\tau_j, \tau_j+\varepsilon]$, for some $\varepsilon > 0$. We denote by (X_{0-}, V_{0-}) the initial price and portfolio values. Then, the number of stocks in the portfolio associated to a strategy $(a,b,\nu) \in \mathcal{A}_k \times \mathcal{U}_k$ is given by

$$Y^\varepsilon = Y + \sum_{j=1}^{k} \mathbf{1}_{[\tau_j,T]} \left[-\delta_j + \varepsilon^{-1}\delta_j(\cdot \wedge (\tau_j+\varepsilon) - \tau_j)\right], \tag{1.20}$$

and the corresponding stock price and portfolio value dynamics are

$$X^\varepsilon = X_{0-} + \int_0^{\cdot} \sigma(X_s^\varepsilon)dW_s + \int_0^{\cdot} f(X_s^\varepsilon)dY_s^\varepsilon + \int_0^{\cdot} (\mu(X_s^\varepsilon) + a_s(\sigma f')(X_s^\varepsilon))ds \tag{1.21}$$

$$V^\varepsilon = V_{0-} + \int_0^{\cdot} Y_s^\varepsilon dX_s^\varepsilon + \frac{1}{2}\int_0^{\cdot} a_s^2 f(X_s^\varepsilon)ds. \tag{1.22}$$

When passing to the limit $\varepsilon \to 0$, we obtain the convergence of $Z^\varepsilon := (X^\varepsilon, Y^\varepsilon, V^\varepsilon)$ to $Z = (X, Y, V)$ with (X, V) defined in (1.23)-(1.24) below. In the following, we only state the convergence of the terminal values, see the proof for a more complete description. It uses the curve x defined in (1.3) above, recall also (1.2).

Proposition 1.2. *Given $(a, b, \nu) \in \mathcal{A} \times \mathcal{U}$, let $Z = (X, Y, V)$ be defined by* (1.19) *and*

$$
\begin{aligned}
X &= X_{0-} + \int_0^{\cdot} \sigma(X_s)dW_s + \int_0^{\cdot} f(X_s)dY_s^c + \int_0^{\cdot} (\mu(X_s) + a_s(\sigma f')(X_s))ds \\
&\quad + \int_0^{\cdot} \int \Delta\mathrm{x}(X_{s-}, \delta)\nu(d\delta, ds) \qquad (1.23) \\
V &= V_{0-} + \int_0^{\cdot} Y_s dX_s^c + \frac{1}{2}\int_0^{\cdot} a_s^2 f(X_s)ds \\
&\quad + \int_0^{\cdot} \int (Y_{s-}\Delta\mathrm{x}(X_{s-}, \delta) + \mathfrak{I}(X_{s-}, \delta))\, \nu(d\delta, ds) \qquad (1.24)
\end{aligned}
$$

where $X^c := X - \int_0^{\cdot} \int \Delta\mathrm{x}(X_{s-}, \delta)\nu(d\delta, ds)$ and

$$
\mathfrak{I}(x, z) := \int_0^z s f(\mathrm{x}(x, s))ds, \quad \text{for } x, z \in \mathbb{R}. \qquad (1.25)
$$

Set $Z^\varepsilon := (X^\varepsilon, V^\varepsilon, Y^\varepsilon)$. Then, there exists a constant $C > 0$ such that

$$
\mathbb{E}\left[|Z^\varepsilon_{T+\varepsilon} - Z_T|^2\right] \leq C(\varepsilon + \mathbb{P}[\sup_{t \leq T} \nu(\mathbb{R}, [t, t+\varepsilon]) \geq 2]^{\frac{1}{2}}),
$$

for all $\varepsilon \in (0, 1)$. Moreover,

$$
\lim_{\varepsilon \to 0} \mathbb{P}[\sup_{t \leq T} \nu(\mathbb{R}, [t, t+\varepsilon]) \geq 2] = 0.
$$

Proof. In all this proof, we denote by C a generic positive constant which does not depend on ε, and may change from line to line. Here again, we shall use repeatedly (H1) and the fact that a and b are bounded by some constant k, in the $dt \times d\mathbb{P}$-a.e. sense.

Let ν be of the form (1.18) for some $k \geq 0$ and note that the last claim simply follows from the fact that $\{\tau_{j+1} - \tau_j \geq \varepsilon\} \uparrow \Omega$ up to a $\mathbb{P}$-null set for all $j \leq k$.
Step 1. We first consider the case where $\tau_{j+1} \geq \tau_j + \varepsilon$ for all $j \geq 1$. Again, the estimate on $|Z^\varepsilon_{T+\varepsilon} - Z_T|$ follows from simple observations and standard estimates, and we only highlight the main ideas. We will indeed prove that for $1 \leq j \leq k+1$

$$
\mathbb{E}\left[\sup_{[\tau_{j-1}+\varepsilon, \tau_j)} |Z - Z^\varepsilon|^2 + \sup_{0 \leq s \leq \varepsilon} \mathbb{E}[|Z_{\tau_j + s} - Z^\varepsilon_{\tau_j+\varepsilon}|^2\right] \leq C\varepsilon, \qquad (1.26)
$$

where we use the convention $\tau_0 = 0$ and $\tau_{k+1} = T$. The result is trivial for (Y^ε, Y) since they are equal on each interval $[\tau_{j-1} + \varepsilon, \tau_j)$ and (a, b) is bounded.

a. We first prove a stronger result for (X^ε, X). Fix $p \in \{2,4\}$. Let x^ε be the solution of the ordinary differential equation

$$\mathrm{x}^\varepsilon_t = X_{\tau_j-} + \int_0^t \frac{\delta_j}{\varepsilon} f(\mathrm{x}^\varepsilon_s) ds.$$

Set $\Delta X^\varepsilon := X^\varepsilon - \mathrm{x}^\varepsilon_{\cdot-\tau_j}$. Itô's Lemma leads to

$$\begin{aligned} d(\Delta X^\varepsilon_t)^p &= p(\Delta X^\varepsilon_t)^{p-1}\alpha^{1,\varepsilon}_t dt + \frac{p(p-1)}{2}(\Delta X^\varepsilon_t)^{p-2}(\alpha^{2,\varepsilon}_t)^2 dt \\ &\quad + p(\Delta X^\varepsilon_t)^{p-1}\alpha^{2,\varepsilon}_t dW_t \\ &\quad + p\frac{\delta_j}{\varepsilon}(\Delta X^\varepsilon_t)^{p-1}(f(X^\varepsilon_t) - f(\mathrm{x}^\varepsilon_{t-\tau_j}))dt \end{aligned}$$

on $[\tau_j, \tau_j+\varepsilon]$, in which $\alpha^{1,\varepsilon}$ and $\alpha^{2,\varepsilon}$ are bounded processes. The inequality $x^{p-1} \le x^{p-2} + x^p$, the Lipschitz continuity of f and Gronwall's Lemma then imply

$$\sup_{0\le t\le\varepsilon} \mathbb{E}\left[|X^\varepsilon_{\tau_j+t} - \mathrm{x}^\varepsilon_t|^p\right] \le C\mathbb{E}\left[|X^\varepsilon_{\tau_j-} - X_{\tau_j-}|^p + \int_0^\varepsilon |X^\varepsilon_{\tau_j+s} - \mathrm{x}^\varepsilon_s|^{p-2} ds\right].$$

We now use a simple change of variables to obtain

$$\mathrm{x}^\varepsilon_\varepsilon = \mathrm{x}(X_{\tau_j-}, \delta_j) = X_{\tau_j},$$

in which x is defined in (1.3), while

$$\sup_{0\le t\le\varepsilon} \mathbb{E}\left[|X_{\tau_j+t} - X_{\tau_j}|^p\right] \le C\varepsilon^{\frac{p}{2}}.$$

Since X and X^ε have the same dynamics on $[\tau_j+\varepsilon, \tau_{j+1})$, this shows that

$$\begin{aligned} \mathbb{E}\left[\sup_{[\tau_j+\varepsilon,\tau_{j+1})} |X_t - X^\varepsilon_t|^p\right] &\le C\mathbb{E}\left[|X_{\tau_j+\varepsilon} - X^\varepsilon_{\tau_j+\varepsilon}|^p\right] \\ &\le C\mathbb{E}\left[|\mathrm{x}^\varepsilon_\varepsilon - X^\varepsilon_{\tau_j+\varepsilon}|^p + |X_{\tau_j+\varepsilon} - X_{\tau_j}|^p\right] \\ &\le C\mathbb{E}\left[|X^\varepsilon_{\tau_j-} - X_{\tau_j-}|^p + \int_0^\varepsilon |X^\varepsilon_{\tau_j+s} - \mathrm{x}^\varepsilon_s|^{p-2} ds + \varepsilon^{\frac{p}{2}}\right]. \end{aligned}$$

For $p = 2$, this provides

$$\mathbb{E}\left[\sup_{[\tau_{j-1}+\varepsilon,\tau_j)} |X - X^\varepsilon|^p + \sup_{0\le s\le\varepsilon} \mathbb{E}[|X_{\tau_j+s} - X^\varepsilon_{\tau_j+\varepsilon}|^p\right] \le C\varepsilon^{\frac{p}{2}},$$

by induction over j, and the case $p=4$ then follows from the above. For later use, note that the estimate

$$\sup_{0\le t\le\varepsilon} \mathbb{E}\left[|X^\varepsilon_{\tau_j+t}-\mathrm{x}^\varepsilon_t|^4\right] \le C\varepsilon^2 \tag{1.27}$$

is a by-product of our analysis.

b. The estimate on $V-V^\varepsilon$ is proved similarly. We introduce

$$\mathrm{v}^\varepsilon_t := V_{\tau_j-} + \int_0^t \frac{\delta_j^2}{\varepsilon^2} s f(\mathrm{x}^\varepsilon_s)ds + Y_{\tau_j-}\int_0^t \frac{\delta_j}{\varepsilon} f(\mathrm{x}^\varepsilon_s)ds = V_{\tau_j-} + \int_0^t Y^\varepsilon_s \frac{\delta_j}{\varepsilon} f(\mathrm{x}^\varepsilon_s)ds,$$

and obtain a first estimate by using (1.27):

$$\begin{aligned}\mathbb{E}\left[|V^\varepsilon_{\tau_j+t}-\mathrm{v}^\varepsilon_t|^2\right] &\le C\mathbb{E}\left[|V^\varepsilon_{\tau_j-}-V_{\tau_j-}|^2+\varepsilon+\left(\int_0^\varepsilon \varepsilon^{-1}Y^\varepsilon_{\tau_j+s}\delta_j|X^\varepsilon_{\tau_j+s}-\mathrm{x}^\varepsilon_s|ds\right)^2\right]\\ &\le C\mathbb{E}\left[|V^\varepsilon_{\tau_j-}-V_{\tau_j-}|^2+\varepsilon\right],\end{aligned}$$

for $0\le t\le\varepsilon$. Then, we observe that

$$\mathrm{v}^\varepsilon_\varepsilon = V_{\tau_j-} + \mathfrak{I}(X_{\tau_j-},\delta_j) + Y_{\tau_j-}\Delta\mathrm{x}(X_{\tau_j-},\delta_j) = V_{\tau_j},$$

while

$$\sup_{0\le t\le\varepsilon}\mathbb{E}\left[|V_{\tau_j+t}-V_{\tau_j}|^2\right]\le C\varepsilon.$$

By using the estimate on $X-X^\varepsilon$ obtained in a., we then show that

$$\mathbb{E}\left[\sup_{[\tau_j+\varepsilon,\tau_{j+1})}|V_t-V^\varepsilon_t|^2\right] \;\le\; C\mathbb{E}\left[|V_{\tau_j+\varepsilon}-V^\varepsilon_{\tau_j+\varepsilon}|^2+\varepsilon\right],$$

and conclude by using an induction over j.

Step 2. We now consider the general case. We define

$$\tau^\varepsilon_{j+1} = (\varepsilon+\tau^\varepsilon_j)\vee\tau_{j+1}\ ,\ \delta^\varepsilon_{j+1} = \int_{(\tau^\varepsilon_j,\tau^\varepsilon_{j+1}]}\delta\nu(d\delta,dt)\ ,\ j\ge 1,$$

where $(\tau^\varepsilon_1,\delta^\varepsilon_1)=(\tau_1,\delta_1)$. On $E_\varepsilon := \{\min_{j\le k-1}(\tau_{j+1}-\tau_j)\ge\varepsilon\}$, $(\tau^\varepsilon_j,\delta^\varepsilon_j)_{j\ge1} = (\tau_j,\delta_j)_{j\ge1}$. Hence, it follows from Step 1. that

$$\mathbb{E}\left[|Z^\varepsilon_{T+\varepsilon}-Z_T|^2\right]\le C\varepsilon + C\mathbb{E}\left[|\tilde{Z}^\varepsilon_{T+\varepsilon}|^4+|Z_T|^4\right]^{\frac12}\mathbb{P}[E^c_\varepsilon]^{\frac12},$$

in which $\tilde{Z}^\varepsilon$ stands for the dynamics associated to $(\tau^\varepsilon_j,\delta^\varepsilon_j)_{j\ge1}$. It now follows from standard estimates that $(\tilde{Z}^\varepsilon_{T+\varepsilon})_{0<\varepsilon\le1}$ and Z_T are bounded in $\mathbf{L}^4$. □

We conclude this section with a proposition collecting some important properties of the functions x and $\mathfrak{I}$ which appear in Proposition 1.1. They will be used in the subsequent section.

Proposition 1.3. *For all $x, y, \iota \in \mathbb{R}$,*

(i) $\mathrm{x}(\mathrm{x}(x,\iota), -y-\iota) = \mathrm{x}(x,-y)$,

(ii) $f(x)\partial_x \mathrm{x}(x,y) = \partial_y \mathrm{x}(x,y) = f(\mathrm{x}(x,y))$,

(iii) $\mathfrak{I}(\mathrm{x}(\mathrm{x}(x,\iota),-y-\iota), y+\iota) - \mathfrak{I}(\mathrm{x}(x,-y),y) = y\Delta\mathrm{x}(x,\iota) + \mathfrak{I}(x,\iota)$,

(iv) $f(x)\partial_x \mathfrak{I}(x,y) + \Delta\mathrm{x}(x,y) = \partial_y \mathfrak{I}(x,y) = yf(\mathrm{x}(x,y))$.

Proof. (i) is an immediate consequence of the Lipschitz continuity of the function f, which ensures uniqueness of the ODE defining x in (1.3). More generally, it has the flow property, which we shall use in the following arguments. The assertion (ii) is an immediate consequence of the definition of x: $\mathrm{x}(\mathrm{x}(x,\iota), y-\iota) = \mathrm{x}(x,y)$ for $\iota > 0$ and $\partial_y \mathrm{x}(x,0) = f(x)$, so that differentiating at $\iota = 0$ provides (ii). The identity in (iii) follows from direct computations. As for (iv), it suffices to write that $\mathfrak{I}(\mathrm{x}(x,\iota), y-\iota) = \int_\iota^y (t-\iota) f(\mathrm{x}(x,t))dt$ for $\iota > 0$, and again to differentiate at $\iota = 0$. □

Remark 1.3. *It follows from Proposition 1.3 that our model allows round trips at (exactly) zero cost. Namely, if x is the current stock price, v the wealth, and y the number of shares in the portfolio, then performing an immediate jump of size δ makes (x,y,v) jump to $(\mathrm{x}(x,\delta),\ y+\delta,\ v+y\Delta\mathrm{x}(x,\delta)+\mathfrak{I}(x,\delta))$. Passing immediately the opposite order, we come back to the position $(\mathrm{x}(\mathrm{x}(x,\delta),-\delta),\ y+\delta-\delta,\ v+y\Delta\mathrm{x}(x,\delta)+\mathfrak{I}(x,\delta)+(y+\delta)\Delta\mathrm{x}(\mathrm{x}(x,\delta),-\delta)+\mathfrak{I}(\mathrm{x}(x,\delta),-\delta)) = (x,y,v)$, by Proposition 1.3(i)-(iii). This is a desirable property if one wants to have a chance to hedge options perfectly, or more generally to obtain a non-degenerated super-hedging price.*

2 Super-hedging of a European claim

We now turn to the super-hedging problem. From now on, we define the admissible strategies as the Itô processes of the form

$$Y = y + \int_0^\cdot b_s ds + \int_0^\cdot a_s dW_s + \int_0^\cdot \int \delta\nu(d\delta, ds) \tag{2.1}$$

in which $y \in \mathbb{R}$, $(a,b,\nu) \in \mathcal{A} \times \mathcal{U}$ and Y is essentially bounded. If $|Y| \le k$ and $(a,b,\nu) \in \mathcal{A}_k \times \mathcal{U}_k$, then we say that $(a,b,\nu) \in \Gamma_k$, $k \ge 1$, and we let

$$\Gamma := \cup_{k\ge1}\Gamma_k.$$

We will comment in Remark 2.1 below the reason why we restrict to bounded controls.

Given $(t, z) \in \mathrm{D} := [0, T] \times \mathbb{R} \times \mathbb{R} \times \mathbb{R}$, we define

$$Z^{t,z,\gamma} := (X^{t,z,\gamma}, Y^{t,z,\gamma}, V^{t,z,\gamma})$$

as the solution of (1.23)-(2.1)-(1.24) on $[t, T]$ associated to $\gamma \in \Gamma$ and with initial condition $Z^{t,z,\gamma}_{t-} = z$.

2.1 Super-hedging price

A European contingent claim is defined by its payoff function, a measurable map $x \in \mathbb{R} \mapsto (g_0, g_1)(x) \in \mathbb{R}^2$. The first component is the cash-settlement part, i.e. the amount of cash paid at maturity, while g_1 is the delivery part, i.e. the number of units of stocks to be delivered.

An admissible strategy $\gamma \in \Gamma$ allows to super-hedge the claim associated to the payoff g, starting from the initial conditions z at time t if

$$Z^{t,z,\gamma}_T \in \mathrm{G}$$

where

$$\mathrm{G} := \{(x, y, v) \in \mathbb{R} \times \mathbb{R} \times \mathbb{R} \ : \ v - yx \geq g_0(x) \text{ and } y = g_1(x)\}. \tag{2.2}$$

Recall that V stands for the frictionless liquidation value of the portfolio, it is the sum of the cash component and the value YX of the stocks held without taking the liquidation impact into account.

We set

$$\mathcal{G}_k(t, z) := \{\gamma \in \Gamma_k \ : \ Z^{t,z,\gamma}_T \in \mathrm{G}\}\ ,\ \mathcal{G}(t, z) := \cup_{k\geq 1}\mathcal{G}_k(t, z),$$

and define the super-hedging price as

$$w(t, x) := \inf_{k\geq 1} w_k(t, x) \ \text{ where } \ w_k(t, x) := \inf\{v : \mathcal{G}_k(t, x, 0, v) \neq \emptyset\}.$$

For later use, let us make precise what are the T-values of these functions.

Proposition 2.1. *Define*

$$G_k(x) := \inf\{y\mathrm{x}(x, y) + g_0(\mathrm{x}(x, y)) - \mathfrak{I}(x, y) : |y| \leq k \ s.t.\ y = g_1(\mathrm{x}(x, y))\},\ x \in \mathbb{R},$$

and $G := \inf_{k\geq 1} G_k$. *Then,*

$$w_k(T, \cdot) = G_k \quad \textit{and} \quad w(T, \cdot) = G. \tag{2.3}$$

Proof. Set $z=(x,0,v)$ and fix $\gamma=(a,b,\nu)\in\Gamma$. By (1.23)-(1.24), we have

$$Z_T^{T,z,\gamma}=(\mathrm{x}(x,y),y,v+\Im(x,y)) \text{ with } y:=\int\delta\nu(d\delta,\{T\}).$$

In view of (2.2), $Z_T^{T,z,\gamma}\in\mathrm{G}$ is then equivalent to

$$v+\Im(x,y)-y\mathrm{x}(x,y)\geq g_0(\mathrm{x}(x,y)) \text{ and } y=g_1(\mathrm{x}(x,y)).$$

By definition of w (resp. w_k), we have to compute the minimal v for which this holds for some $y\in\mathbb{R}$ (resp. $|y|\leq k$). □

Remark 2.1. *Let us conclude this section with a comment on our choice of the set of bounded controls* Γ*.*

a. First, this ensures that the dynamics of X,Y *and* V *are well-defined. This could obviously be relaxed by imposing* $\mathbf{L}^2_\lambda$ *bounds. However, note that the bound should anyway be uniform. This is crucial to ensure that the dynamic programming principle stated in Section 2.2 is valid, as it uses measurable selection arguments:* $\omega\mapsto\vartheta[\omega]\in\mathbf{L}_2^\lambda$ *does not imply* $\mathbb{E}\left[\|\vartheta[\cdot]\|_{\mathbf{L}_2^\lambda}\right]<\infty$*. See Remark 2.2 below for a related discussion.*

b. In the proof of Theorem 2.1, we will need to perform a change of measure associated to a martingale of the form $dM=-M\chi^a dW$ *in which* χ^a *may explode at a speed* a^2 *if* a *is not bounded. See Step 1. of the proof of Theorem 2.1. In order to ensure that this local martingale is well-defined, and is actually a martingale, one should impose very strong integrability conditions on* a*.*

In order to simplify the presentation, we therefore stick to bounded controls. Many other choices are possible. Note however that, in the case $f\equiv 0$*, a large class of options leads to hedging strategies in our set* Γ*, up to a slight payoff smoothing to avoid the explosion of the delta or the gamma at maturity. This implies that, although the perfect hedging strategy may not belong to* Γ*, at least it is a limit of elements of* Γ *and the super-hedging prices coincide.*

2.2 Dynamic programming

Our control problem is a stochastic target problem as studied in [18]. The aim of this section is to show that it satisfies a version of their geometric dynamic programming principle.

However, the value function w is not amenable to dynamic programming per se. The reason is that it assumes a zero initial stock holding at time t, while the position Y_θ will (in general) not be zero at a later time θ. It is therefore a priori not possible to compare the later wealth process V_θ with the corresponding super-hedging price $w(\theta,X_\theta)$.

Still, a version of the geometric dynamic programming principle can be obtained if we introduce the process

$$\hat{X}^{t,z,\gamma} := \mathrm{x}(X^{t,z,\gamma}, -Y^{t,z,\gamma}) \tag{2.4}$$

which represents the value of the stock immediately after liquidating the stock position.

We refer to Remark 2.2 below for the reason why part (ii) of the following dynamic programming principle is stated in terms of $(w_k)_{k\geq 1}$ instead of w.

Proposition 2.2 (GDP)**.** *Fix $(t,x,v) \in [0,T]\times\mathbb{R}\times\mathbb{R}$.*

(i) *If $v > w(t,x)$ then there exists $\gamma \in \Gamma$ and $y \in \mathbb{R}$ such that*

$$V_\theta^{t,z,\gamma} \quad \geq \quad w(\theta, \hat{X}_\theta^{t,z,\gamma}) + \mathfrak{I}(\hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma}),$$

for all stopping time $\theta \geq t$, where $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$.

(ii) *Fix $k \geq 1$. If $v < w_{2k+2}(t,x)$ then we can not find $\gamma \in \Gamma_k$, $y \in [-k,k]$ and a stopping time $\theta \geq t$ such that*

$$V_\theta^{t,z,\gamma} > w_k(\theta, \hat{X}_\theta^{t,z,\gamma}) + \mathfrak{I}(\hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$$

with $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$.

Proof. Step 1. In order to transform our stochastic target problem into a time consistent one, we introduce the auxiliary value function corresponding to an initial holding y in stocks:

$$\hat{w}(t,x,y) := \inf_{k\geq 1} \hat{w}_k(t,x,y) \ \text{ where } \ \hat{w}_k(t,x,y) := \inf\{v : \mathcal{G}_k(t,x,y,v) \neq \emptyset\}.$$

Note that $w_{k+1}(t,x) \leq \inf\{v : \exists\, y \in [-k,k] \text{ s.t. } \mathcal{G}_k(t, \mathrm{x}(x,y), y, v + \mathfrak{I}(x,y)) \neq \emptyset\}$. This follows from (1.23)-(1.24). Since $\mathrm{x}(\mathrm{x}(x,-y),y) = x$, see Proposition 1.3, this implies that

$$\hat{w}_k(t,x,y) \geq w_{k+1}(t, \mathrm{x}(x,-y)) + \mathfrak{I}(\mathrm{x}(x,-y),y), \tag{2.5}$$

for $|y| \leq k$. Similarly, since $\mathfrak{I}(x,-y) + y\Delta\mathrm{x}(x,-y) = -\mathfrak{I}(\mathrm{x}(x,-y),y)$ by Proposition 1.3, we have

$$\hat{w}_{k+1}(t,x,y) \leq w_k(t, \mathrm{x}(x,-y)) + \mathfrak{I}(\mathrm{x}(x,-y),y). \tag{2.6}$$

Step 2. a. Assume that $v > w(t,x)$. The definition of w implies that we can find $y \in \mathbb{R}$ and $\gamma \in \mathcal{G}(t,z)$ where $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$. By the arguments of

[18, Step 1 proof of Theorem 3.1], $V_\theta^{t,z,\gamma} \geq \hat{w}(\theta, X_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$, for all stopping time $\theta \geq t$. Then, (2.5) applied for $k \to \infty$ provides (i).

b. Assume now that we can find $\gamma \in \Gamma_k$, $y \in [-k, k]$ and a stopping time $\theta \geq t$ such that $V_\theta^{t,z,\gamma} > (w_k + \mathfrak{I})(\theta, \hat{X}_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$, where $z := (\mathrm{x}(x,y), y, v + \mathfrak{I}(x,y))$. By (2.4)-(2.6), $V_\theta^{t,z,\gamma} > \hat{w}_{k+1}(\theta, X_\theta^{t,z,\gamma}, Y_\theta^{t,z,\gamma})$, and it follows from [18, Step 2 proof of Theorem 3.1] and Corollary A.1 that $v + \mathfrak{I}(x,y) \geq \hat{w}_{2k+1}(t, \mathrm{x}(x,y), y)$. We conclude that (ii) holds by appealing to (2.5) and the identities $\mathrm{x}(\mathrm{x}(x,y), -y) = x$ and $\mathfrak{I}(\mathrm{x}(\mathrm{x}(x,y), -y), y) = \mathfrak{I}(x,y)$, see Proposition 1.3. $\square$

We conclude this section with purely technical considerations that justify the form of the above dynamic programming principle. They are of no use for the later developments but may help to clarify our approach.

Remark 2.2. *Part* (ii) *of Proposition 2.2 can not be stated in terms of w. The reason is that measurable selection technics can not be used with the set Γ. Indeed, if $\omega \mapsto \gamma[\omega] \in \Gamma$, then the corresponding bounds depend on ω and are not uniform: a measurable family of controls $\{\gamma[\omega], \omega \in \Omega\}$ does not permit to construct an element in Γ. Part* (i) *of Proposition 2.2 only requires to use a conditioning argument, which can be done within Γ.*

Remark 2.3. *A version of the geometric dynamic programming principle also holds for $(\hat{w}_k)_{k\geq 1}$, this is a by-product of the above proof. It is therefore tempting to try to derive a pde for the function $\hat{w}$. However, the fact that the control b appears linearly in the dynamics of (X, Y, V) makes this problem highly singular, and "standard approaches" do not seem to work. We shall see in Lemma 2.1 that this singularity disappears in the parameterization* $\mathrm{x}(X, -Y)$ *used in Proposition 2.2. Moreover, hedging implies a control on the diffusion part of the dynamics which translates into a strong relation between Y and the space gradient $D\hat{w}(\cdot, X, Y)$. This would lead to a pde set on a curve on the coordinates (t, x, y) depending on $D\hat{w}$ (the solution of the pde).*

2.3 Pricing equation

In order to understand what is the partial differential equation that w should solve, let us state the following key lemma. Although the control b appears linearly in the dynamics of (X, Y, V), the following shows that the singularity this may create does indeed not appear when applying Itô's Lemma to $V - (\varphi + \mathfrak{I})(\cdot, \hat{X}, Y)$, recall (2.4), it is absorbed by the functions x and $\mathfrak{I}$ (compare with Remark 2.3). The proof of this Lemma is postponed to Section 2.5.

Lemma 2.1. *Fix $(t,x,y,v) \in \mathrm{D}$, $z := (x,y,v)$, $\gamma = (a,b,\nu) \in \Gamma$. Then,*

$$\begin{aligned} \hat{X}^{t,z,\gamma} &= \mathrm{x}(x,-y) \\ &+ \int_t^{\cdot} [\hat{\mu}(\hat{X}_s^{t,z,\gamma}, Y_s^{t,z,\gamma}) + (\partial_x \mathrm{x}\mu - \frac{1}{2}\partial_x \mathrm{x} a_s^2 f f')(X_s^{t,z,\gamma}, -Y_s^{t,z,\gamma})]ds \\ &+ \int_t^{\cdot} \hat{\sigma}(\hat{X}_s^{t,z,\gamma}, Y_s^{t,z,\gamma}) dW_s. \end{aligned}$$

Given $\varphi \in C_b^\infty$, set $\mathcal{E}^{t,z,\gamma} := V^{t,z,\gamma} - (\varphi + \mathfrak{I})(\cdot, \hat{X}^{t,z,\gamma}, Y^{t,z,\gamma})$. Then,

$$\begin{aligned} \mathcal{E}^{t,z,\gamma} - \mathcal{E}_t^{t,z,\gamma} &= \int_t^{\cdot} [Y_s^{t,z,\gamma} - \check{Y}_s^{t,z,\gamma}](\mu - f' f a_s^2/2)(X_s^{t,z,\gamma}) ds \\ &+ \int_t^{\cdot} [Y_s^{t,z,\gamma} - \check{Y}_s^{t,z,\gamma}] \sigma(X_s^{t,z,\gamma}) dW_s \\ &+ \int_t^{\cdot} \hat{F}\varphi(s, \hat{X}_s^{t,z,\gamma}, Y_s^{t,z,\gamma}) ds \end{aligned}$$

in which

$$\check{Y}^{t,z,\gamma} := Y^{t,z,\gamma} + \frac{\hat{X}^{t,z,\gamma} - X^{t,z,\gamma}}{f(X^{t,z,\gamma})} + \partial_x \varphi(\cdot, \hat{X}^{t,z,\gamma}) \frac{f(\hat{X}^{t,z,\gamma})}{f(X^{t,z,\gamma})}$$

and

$$\hat{F}\varphi := -\partial_t \varphi - \hat{\mu}\partial_x[\varphi + \mathfrak{I}] - \frac{1}{2}\hat{\sigma}^2 \partial_{xx}^2[\varphi + \mathfrak{I}],$$

where for $(x',y') \in \mathbb{R} \times \mathbb{R}$

$$\hat{\mu}(x',y') := \frac{1}{2}[\partial_{xx}^2 \mathrm{x}\sigma^2](\mathrm{x}(x',y'), -y') \quad \textit{and} \quad \hat{\sigma}(x',y') := (\sigma \partial_x \mathrm{x})(\mathrm{x}(x',y'), -y').$$

Let us now appeal to Proposition 2.2 and apply Lemma 2.1 to $\varphi = w$, assuming that w is smooth and that Proposition 2.2(i) is valid even if we start from $v = w(t,x)$, i.e. assuming that the inf in the definition of w is a min. With the notations of the above lemma, Proposition 2.2(i) formally applied to $\theta = t+$ leads to

$$\begin{aligned} 0 &\leq d\mathcal{E}_t^{t,z,\gamma} \\ &= (y - \hat{y}) \left\{ [\mu - f f' a_t^2/2)(\mathrm{x}(x,y))]dt + \sigma(\mathrm{x}(x,y)) dW_t \right\} \\ &+ \hat{F}w(t, \hat{x}, y) dt \end{aligned}$$

in which

$$\hat{y} = y + \frac{\hat{x} - \mathrm{x}(x,y)}{f(\mathrm{x}(x,y))} + \partial_x w(t,\hat{x}) \frac{f(\hat{x})}{f(\mathrm{x}(x,y))} \quad \text{and} \quad \hat{x} = \mathrm{x}(\mathrm{x}(x,y), -y) = x.$$

Remaining at a formal level, this inequality cannot hold unless $y = \hat{y}$, because $\sigma \neq 0$, and

$$\hat{F}w(t,x,\hat{y}) = \hat{F}w(t,\hat{x},y) \geq 0.$$

This means that w should be a super-solution of

$$F\varphi(t,x) \quad := \quad \hat{F}\varphi(t,x,\hat{y}[\varphi](t,x)) = 0 \tag{2.7}$$

where, for a smooth function φ,

$$\hat{y}[\varphi](t,x) := \mathrm{x}^{-1}(x, x + f(x)\partial_x\varphi(t,x))$$

and x^{-1} denotes the inverse of $\mathrm{x}(x,\cdot)$.

From (ii) of Proposition 2.2, we can actually (formally) deduce that the above inequality should be an equality, and therefore that w should solve (2.7).

In order to give a sense to the above, we assume that

$$\left\{ \begin{array}{c} \mathrm{x}(x,\cdot) \text{ is invertible for all } x \in \mathbb{R} \\ (x,z) \in \mathbb{R}\times\mathbb{R} \mapsto \mathrm{x}^{-1}(x,z) \;\text{ is } C^2. \end{array} \right. \tag{H2}$$

In view of (2.3), we therefore expect w to be a solution of

$$F\varphi\mathbf{1}_{[0,T[} + (\varphi - G)\mathbf{1}_{\{T\}} = 0 \;\text{ on } [0,T]\times\mathbb{R}. \tag{2.8}$$

Since w may not be smooth and (ii) of Proposition 2.2 is stated in terms of w_k instead of w, we need to consider the notion of viscosity solutions and the relaxed semi-limits of $(w_k)_{k\geq 1}$. We therefore define

$$w_*(t,x) := \liminf_{(t',x',k)\to(t,x,\infty)} w_k(t',x') \text{ and } w^*(t,x) := \limsup_{(t',x',k)\to(t,x,\infty)} w_k(t',x'),$$

in which the limits are taken over $t' < T$, as usual. Note that w_* actually coincides with the lower-semicontinuous envelope of w, this comes from the fact that $w = \inf_{k\geq 1} w_k = \lim_{k\to\infty} \downarrow w_k$, by construction.

We are now in position to state the main result of this section. In the following, we assume that

$$\left\{ \begin{array}{c} G \;\text{ is continuous and }\; G_k \downarrow G \;\text{ uniformly on compact sets.} \\ w_* \text{ and } w^* \text{ are finite on } [0,T]\times\mathbb{R}. \end{array} \right. \tag{H3}$$

The first part of (H3) will be used to obtain the boundary condition. The second part is natural since otherwise our problem would be ill-posed.

Theorem 2.1 (Pricing equation)**.** *The functions w_* and w^* are respectively a viscosity super- and a subsolution of* (2.8)*. If they are bounded and* $\inf f > 0$*, then $w = w_* = w^*$ and w is the unique bounded viscosity solution of* (2.8)*. If in addition G is bounded and C^2 with G, G', G'' Hölder continuous, then $w \in C^{1,2}([0,T) \times \mathbb{R}) \cap C^0([0,T] \times \mathbb{R})$.*

The proof is reported in Section 2.5. Let us now discuss the verification counterpart.

Remark 2.4 (Verification)**.** *Assume that φ is a smooth solution of* (2.8) *and that we can find $(a,b) \in \mathcal{A}$ such that the following system holds on $[t,T)$:*

$$
\begin{aligned}
X &= x + \Delta\mathrm{x}(x, \hat{y}[\varphi](t,x)) + \int_t^{\cdot} \sigma(X_s) dW_s + \int_0^{\cdot} f(X_s) dY_s^c \\
&\quad + \int_0^{\cdot} (\mu(X_s) + a_s(\sigma f')(X_s)) ds + \Delta\mathrm{x}(X_{T-}, -Y_{T-}) \mathbf{1}_{\{T\}} \\
Y &= \hat{y}[\varphi](t,x) + \int_t^{\cdot} b_s ds + \int_t^{\cdot} a_s dW_s - Y_{T-} \mathbf{1}_{\{T\}} \\
&= \mathrm{x}^{-1}(\hat{X}, \hat{X} + (f\partial_x\varphi)(\cdot, \hat{X})) - Y_{T-} \mathbf{1}_{\{T\}} \\
\hat{X} &:= \mathrm{x}(X, -Y) \\
V &= \varphi(t,x) + \mathfrak{I}(x, \hat{y}[\varphi](t,x)) + \int_t^{\cdot} Y_s dX_s^c + \frac{1}{2} \int_0^{\cdot} a_s^2 f(X_s) ds \\
&\quad + (Y_{T-} \Delta\mathrm{x}(X_{T-}, -Y_{T-}) + \mathfrak{I}(X_{T-}, -Y_{T-})) \mathbf{1}_{\{T\}}.
\end{aligned}
$$

a. Note that $\hat{X}_t = \mathrm{x}(X_t, -Y_t) = \mathrm{x}(\mathrm{x}(x, \hat{y}[\varphi](t,x)), -\hat{y}[\varphi](t,x)) = x$, recall Proposition 1.3(i), so that $Y_t = \hat{y}[\varphi](t,x) = \mathrm{x}^{-1}(\hat{X}_t, \hat{X}_t + (f\partial_x\varphi)(t, \hat{X}_t))$. We therefore need to find (a,b) such that $X = \mathrm{x}(\hat{X}, Y) = \hat{X} + (f\partial_x\varphi)(\cdot, \hat{X})$. This amounts to solving:

$$
\begin{aligned}
\sigma(X) + f(X)a &= \hat{\sigma}(\hat{X}, Y) \partial_x \psi(\cdot, \hat{X}) \\
f(X)b + (\mu + a\sigma f')(X) &= (\hat{\mu}(\hat{X}, Y) + (\partial_x \mathrm{x} \mu - \frac{1}{2} \partial_x \mathrm{x} a_s^2 f f')(X, -Y)) \partial_x \psi(\cdot, X) \\
&\quad + \frac{1}{2} \hat{\sigma}^2(\hat{X}, Y) \partial_{xx}^2 \psi(\cdot, \hat{X})
\end{aligned}
$$

where $\psi(t,x) := x + (f\partial_x\varphi)(t,x)$. Since $f > 0$, this system has a solution. Under additional smoothness and boundedness assumption, $(a,b) \in \mathcal{A}$.

b. Let $\check{Y}$ be as in Lemma 2.1 for the above dynamics. Since $X = \mathrm{x}(\hat{X}, Y) = \hat{X} + (f\partial_x\varphi)(\cdot, \hat{X})$ on $[t,T)$ by construction, we have $\check{Y} = Y$ on $[t,T)$. Then, it follows from Lemma 2.1 and (2.7)-(2.8) *that*

$$
V_{T-} = \varphi(T, \hat{X}_{T-}) + \mathfrak{I}(\hat{X}_{T-}, Y_{T-}) = G(\hat{X}_{T-}) + \mathfrak{I}(\hat{X}_{T-}, Y_{T-}).
$$

Since $X_T = \hat{X}_{T-}$ and $Y_{T-}\Delta\mathrm{x}(X_{T-}, -Y_{T-}) + \mathfrak{I}(X_{T-}, -Y_{T-}) + \mathfrak{I}(\hat{X}_{T-}, Y_{T-}) = 0$, see Proposition 1.3, this implies that $V_T = G(X_T)$. Hence, the hedging strategy consists in taking an initial position is stocks equal to $Y_t = \hat{y}[\varphi](t,x)$ and then to use the control (a,b) up to T. A final immediate trade is performed at T. In particular, the number of stocks Y is continuous on (t,T).

2.4 An example: the fixed impact case

In this section, we consider the simple case of a constant impact function f: $f(x) = \lambda > 0$ for all $x \in \mathbb{R}$. This is certainly a too simple model, but this allows us to highlight the structure of our result as the pde simplifies in this case. Indeed, for

$$\mathrm{x}(x,y) = x + y\lambda \quad \text{and} \quad \mathfrak{I}(x,y) = \frac{1}{2}y^2\lambda,$$

we have

$$\hat{\mu}(x,y) = 0 \ \ , \ \hat{\sigma}(x,y) := \sigma(x+y\lambda) \ \ , \ \hat{y}[\varphi] := \partial_x\varphi.$$

The pricing equation is given by a local volatility model in which the volatility depends on the hedging price itself, and therefore on the claim (g_0, g_1) to be hedged:

$$0 = -\partial_t\varphi(t,x) - \frac{1}{2}\sigma^2(x + \partial_x\varphi\lambda)\partial^2_{xx}\varphi(t,x).$$

As for the process Y in the verification argument of Remark 2.4, it is given by

$$Y = \partial_x\varphi(\cdot, \hat{X}) = \partial_x\varphi(\cdot, X - \lambda Y).$$

This shows that the hedging strategy (if it is well-defined) consists in following the usual Δ-hedging strategy but for a $\Delta = \partial_x\varphi$ computed at the value of the stock $\hat{X}$ which would be obtained if the position in stocks was liquidated.

Note that we obtain the usual heat equation when σ is constant. This is expected, showing the limitation of the fixed impact model. To explain this, let us consider the simpler case $g_1 = 0$ and use the notations of Remark 2.4. We also set $\mu = 0$ for ease of notations. Since σ is constant, the strategy Y does not affect the coefficients in the dynamics of X, it just produces a shift λdY each time we buy or sell. Since $Y_T = 0$ (after the final jump), and $Y_{t-} = 0$, the total impact is null:

$X_T = X_{t-} + \sigma(W_T - W_t)$. As for the wealth process, we have

$$
\begin{aligned}
V_T &= \varphi(t,x) + \frac{1}{2}Y_t^2\lambda + \int_t^T Y_s dX_s^c + \frac{1}{2}\int_t^T a_s^2 \lambda ds - Y_{T-}^2\lambda + \frac{1}{2}Y_{T-}^2\lambda \\
&= \varphi(t,x) + \int_t^T Y_s \sigma dW_s + \frac{1}{2}\lambda(Y_t^2 - Y_{T-}^2) + \int_t^T \lambda Y_s dY_s^c + \frac{1}{2}\int_t^T a_s^2\lambda ds \\
&= \varphi(t,x) + \int_t^T Y_s \sigma dW_s.
\end{aligned}
$$

Otherwise stated, the liquidation costs are cancelled: when buying, the trader pays a cost but moves the price up, when selling back, he pays a cost again but sell at a higher price. If there is no effect on the underlying dynamics of X and f is constant, this perfectly cancels.

However, the hedging strategy is still affected: $Y = \partial_x \varphi(\cdot, X - \lambda Y)$ on $[0,T)$.

2.5 Proof of the pde characterization

2.5.1 The key lemma

We first provide the proof of our key result.

Proof of Lemma 2.1. To alleviate the notations, we omit the super-scripts.
a. We first observe from Proposition 1.3(i) that $\mathrm{x}(X,-Y)$ has continuous paths, while Proposition 1.3(ii) implies that $f\partial_x\mathrm{x} - \partial_y\mathrm{x} = 0$ (and therefore $f'\partial_x\mathrm{x} + f\partial^2_{xx}\mathrm{x} - \partial^2_{xy}\mathrm{x} = 0$). Using Itô's Lemma, this leads to

$$
\begin{aligned}
d\mathrm{x}(X_s,-Y_s) \quad &= \quad (\mu - \frac{1}{2}a_s^2 ff')(X_s)\partial_x\mathrm{x}(X_s,-Y_s)ds + \sigma(X_s)\partial_x\mathrm{x}(X_s,-Y_s)dW_s \\
&\quad + \frac{1}{2}\left[\sigma^2\partial^2_{xx}\mathrm{x} - a_s^2 f\partial^2_{xy}\mathrm{x} + a_s^2\partial^2_{yy}\mathrm{x}\right](X_s,-Y_s)ds.
\end{aligned}
$$

We now use the identity $f\partial^2_{xy}\mathrm{x} - \partial^2_{yy}\mathrm{x} = 0$, which also follows from Proposition 1.3(ii), to simplify the above expression into

$$
\begin{aligned}
d\mathrm{x}(X_s,-Y_s) \quad &= \quad [\partial_x\mathrm{x}(\mu - \frac{1}{2}a_s^2 ff') + \frac{1}{2}\partial^2_{xx}\mathrm{x}\sigma^2](X_s,-Y_s)ds \\
&\quad + (\sigma\partial_x\mathrm{x})(X_s,-Y_s)dW_s.
\end{aligned}
$$

b. Similarly, it follows from Proposition 1.3(iii) that $V - \mathfrak{I}(\hat{X}, Y)$ has continuous paths, and so does $\mathcal{E}$ by a. Before applying Itô's lemma to derive the dynamics of $\mathcal{E}$, let us observe that $\partial_y\mathfrak{I}(\mathrm{x}(x,-y),y) = yf(\mathrm{x}(\mathrm{x}(x,-y),y)) = yf(x)$

and that $\partial^2_{yy}\mathfrak{I}(\mathrm{x}(x,-y),y) = y(ff')(x) + f(x)$. Also note that $\hat\sigma(\mathrm{x}(x,-y),y) = \sigma(x)\partial_x \mathrm{x}(x,-y)$. Then, using the dynamics of $\hat X$ derived above, we obtain

$$\begin{aligned} d\mathcal{E}_s =& (Y_s - \check Y_s)\sigma(X_s)dW_s + (Y_s - \check Y_s)[\mu - \frac{1}{2}a_s^2(ff')](X_s)ds + \hat F\varphi(s,\hat X_s, Y_s)ds \\ &+\ a_s\sigma(X_s)[Y_s f'(X_s) - \partial_x \mathrm{x}(X_s,-Y_s)\partial^2_{xy}\mathfrak{I}(\hat X_s,Y_s)]ds, \end{aligned}$$

where

$$\check Y := \partial_x(\varphi+\mathfrak{I})(\cdot,\hat X, Y)\partial_x \mathrm{x}(X,-Y).$$

By Proposition 1.3(ii)(iv), $f(x)\partial^2_{xy}\mathfrak{I}(x,y) = \partial_y[yf(\mathrm{x}(x,y)) - \Delta\mathrm{x}(x,y)] = y(f'f)(\mathrm{x}(x,y))$. Since $\partial_x\mathrm{x}(x,-y) = f(\mathrm{x}(x,-y))/f(x)$, see Proposition 1.3(ii), it follows that

$$\partial_x\mathrm{x}(X,-Y)\partial^2_{xy}\mathfrak{I}(\mathrm{x}(X,-Y),Y) = Yf'(X),$$

which implies

$$d\mathcal{E}_s = (Y_s - \check Y_s)\sigma(X_s)dW_s + (Y_s - \check Y_s)[\mu - \frac{1}{2}a_s^2(ff')](X_s)ds + \hat F\varphi(s,\hat X_s,Y_s)ds.$$

We now deduce from Proposition 1.3 that

$$\begin{aligned} \partial_x\mathfrak{I}(\hat X, Y) &= \frac{-\Delta\mathrm{x}(\hat X,Y) + Yf(\mathrm{x}(\hat X,Y))}{f(\hat X)} = \frac{\hat X - X + Yf(X)}{f(\hat X)} \\ \partial_x\mathrm{x}(X,-Y) &= f(\hat X)/f(X), \end{aligned}$$

so that

$$\check Y = \partial_x\varphi(\cdot,\hat X)\frac{f(\hat X)}{f(X)} + \frac{\hat X - X}{f(X)} + Y.$$

□

2.5.2 Super- and subsolution properties

We now prove the super- and subsolution properties of Theorem 2.1.

Supersolution property. We first prove the supersolution property. It follows from similar arguments as in [5]. Let φ be a C_b^∞ function, and $(t_o,x_o)\in[0,T]\times\mathbb{R}$ be a strict (local) minimum point of $w_* - \varphi$ such that $(w_*-\varphi)(t_o,x_o) = 0$.

a. We first assume that $t_o < T$ and $F\varphi(t_o,x_o) < 0$, and work towards a contradiction. In view of (2.7),

$$\hat F\varphi(t,x,y) < 0 \ \text{ if } (t,x)\in B \text{ and } |y - \hat y[\varphi](t,x)| \le \varepsilon,$$

for some open ball $B \subset [0,T[\times\mathbb{R}$ which contains (t_o,x_o), and some $\varepsilon > 0$. Since x^{-1} is continuous, this implies that

$$\hat{F}\varphi(t,x,y) < 0 \text{ if } (t,x) \in B \text{ and } |x + \partial_x\varphi(t,x)f(x) - \mathrm{x}(x,y)| \leq \varepsilon f(\mathrm{x}(x,y)), \tag{2.9}$$

after possibly changing B and ε. Let $(t_n,x_n)_n$ be a sequence in B that converges to (t_o,x_o) and such that $w(t_n,x_n) \to w_*(t_o,x_o)$ (recall that w_* coïncides with the lower-semicontinuous envelope of w). Set $v_n := w(t_n,x_n) + n^{-1}$. It follows from Proposition 2.2(i) that we can find $(a^n,b^n,\nu^n) = \gamma_n \in \Gamma$ and $y_n \in \mathbb{R}$ such that

$$V_{\theta_n}^{t_n,z_n,\gamma_n} \;\; \geq \;\; w(\theta_n,\hat{X}_{\theta_n}^{t_n,z_n,\gamma_n}) + \Im(\hat{X}_{\theta}^{t_n,z_n,\gamma_n},Y_{\theta_n}^{t_n,z_n,\gamma_n}), \tag{2.10}$$

where $z_n := (\mathrm{x}(x_n,y_n),y_n,v_n + \Im(x_n,y_n))$ and θ_n is the first exit time after t_n of $(\cdot,\hat{X}^{t_n,z_n,\gamma_n})$ from B (note that $\hat{X}_{t_n}^{t_n,z_n,\gamma_n} = \mathrm{x}(\mathrm{x}(x_n,y_n),-y_n) = x_n$). In the following, we use the simplified notations X^n, $\hat{X}^n$, V^n and Y^n for the corresponding quantities indexed by (t_n,z_n,γ_n). Since (t_o,x_o) reaches a strict minimum $w_* - \varphi$, this implies

$$V_{\theta_n}^n \;\; \geq \;\; \varphi(\theta_n,\hat{X}_{\theta_n}^n) + \Im(\hat{X}_{\theta}^n,Y_{\theta_n}^n) + \iota, \tag{2.11}$$

for some $\iota > 0$. Let $\check{Y}^n$ be as in Lemma 2.1 and observe that

$$\check{Y}^n - Y^n = \frac{\hat{X}^n + \partial_x\varphi(\cdot,\hat{X}^n)f(\hat{X}^n) - \mathrm{x}(\hat{X}^n,Y^n)}{f(\mathrm{x}(\hat{X}^n,Y^n))}. \tag{2.12}$$

Set

$$\chi^n := \frac{(\mu - f'f(a_s^n)^2/2)(X^n)}{\sigma(X^n)} + \frac{\hat{F}\varphi(\cdot,\hat{X}^n,Y^n)}{(Y^n - \check{Y}^n)\sigma(X^n)}\mathbf{1}_{|Y^n - \check{Y}^n| \geq \varepsilon}$$

and consider the measure $\mathbb{P}^n$ defined by

$$\frac{d\mathbb{P}_n}{d\mathbb{P}} = M_{\theta_n}^n \text{ where } M^n = 1 - \int_{t_n}^{\cdot\wedge\theta_n} M_s^n\chi_s^n dW_s.$$

Then, it follows from (2.11), Lemma 2.1, (2.9) and (2.12) that

$$\begin{aligned}
\iota &\leq \mathbb{E}^{\mathbb{P}_n}[V_{\theta_n}^n - (\varphi + \Im)(\theta_n,\hat{X}_{\theta_n}^n,Y_{\theta_n}^n)] \\
&\leq v_n + \Im(x_n,y_n) - (\varphi + \Im)(t_n,\mathrm{x}(\mathrm{x}(x_n,y_n),-y_n),y_n) \\
&= v_n - \varphi(t_n,x_n).
\end{aligned}$$

The right-hand side goes to 0, which is the required contradiction.

b. We now explain how to modify the above proof for the case $t_o = T$. After possibly replacing $(t,x) \mapsto \varphi(t,x)$ by $(t,x) \mapsto \varphi(t,x) - \sqrt{T-t}$, we can assume that

$\partial_t\varphi(t,x) \to \infty$ as $t \to T$, uniformly in x on each compact set. Then (2.9) still holds for B of the form $[T-\eta, T) \times B(x_o)$ in which $B(x_o)$ is an open ball around x_o and $\eta > 0$ small. Assume that $\varphi(T, x_o) < G(x_o)$. Then, after possibly changing $B(x_o)$, we have $\varphi(T, \cdot) \le G - \iota_1$ on $B(x_o)$, for some $\iota_1 > 0$. Then, with the notations of a., we deduce from (2.3)-(2.10) that

$$V^n_{\theta_n} \quad \ge \quad \varphi(\theta_n, \hat{X}^n_{\theta_n}) + \Im(\hat{X}^n_\theta, Y^n_{\theta_n}) + \iota_1 \wedge \iota_2,$$

in which $\iota_2 := \min\{(w_* - \varphi)(t,x) : (t,x) \in [t_o - \eta, T) \times \partial B(x_o)\} > 0$ and θ_n is now the minimum between T and the first time after t_n at which $\hat{X}^n$ exists $B(x_o)$. The contradiction is then deduced from the same arguments as above. □

Subsolution property. We now turn to the subsolution property. Again the proof is close to [5], except that we have to account for the specific form of the dynamic programming principle stated in Proposition 2.2(ii). Let φ be a C^∞_b function, and $(t_o, x_o) \in [0,T] \times \mathbb{R}$ be a strict (local) maximum point of $w^* - \varphi$ such that $(w^* - \varphi)(t_o, x_o) = 0$. By [2, Lemma 4.2], we can find a sequence $(k_n, t_n, x_n)_{n\ge1}$ such that $k_n \to \infty$, (t_n, x_n) is a local maximum point of $w^*_{k_n} - \varphi$ and $(t_n, x_n, w_{k_n}(t_n, x_n)) \to (t_o, x_o, w^*(t_o, x_o))$.

a. As above, we first assume that $t_o < T$. Set $\varphi_n(t,x) := \varphi(t,x) + |t - t_n|^2 + |x - x_n|^4$ and assume that $F\varphi(t_o, x_o) > 0$. Then, $F\varphi_n > 0$ on a open neighborhood B of (t_o, x_o) which contains (t_n, x_n), for all n large enough. Since we are going to localize the dynamics, we can modify φ_n, σ, μ and f in such a way that they are identically equal to 0 outside a compact $A \supset B$. It then follows from Remark 2.4 a. that, after possibly changing $n \ge 1$, we can find $(b^n, a^n) \in \mathcal{A}_{k_n}$ such that the following admits a strong solution:

$$\begin{aligned}
X^n &= x_n + \Delta\mathrm{x}(x_n, \hat{y}[\varphi_n](t_n, x_n)) + \int_{t_n}^{\cdot} \sigma(X^n_s)dW_s + \int_{t_n}^{\cdot} f(X^n_s)dY^{n,c}_s \\
&\quad + \int_{t_n}^{\cdot} (\mu(X_s) + a^n_s(\sigma f')(X^n_s))ds \\
Y^n &= \hat{y}[\varphi_n](t_n, x_n) + \int_{t_n}^{\cdot} b^n_s ds + \int_{t_n}^{\cdot} a^n_s dW_s \\
&= \mathrm{x}^{-1}(\hat{X}^n, \hat{X}^n + (f\partial_x\varphi_n)(\cdot, \hat{X}^n)) \\
\hat{X}^n &:= \mathrm{x}(X^n, -Y^n) \\
V^n &= v_n + \Im(x_n, \hat{y}[\varphi_n](t_n, x_n)) + \int_{t_n}^{\cdot} Y^n_s dX^{n,c}_s + \frac{1}{2}\int_{t_n}^{\cdot} (a^n_s)^2 f(X^n_s)ds.
\end{aligned}$$

In the above, we have set $v_n := w_{k_n}(t_n, x_n) - n^{-1}$. Observe that the construction of Y^n ensures that it coincides with the corresponding process $\check{Y}^n$ of Lemma 2.1. Also note that $\hat{X}^n_{t_n} = \mathrm{x}(\mathrm{x}(x_n, y_n), -y_n) = x_n$, and let θ_n be the first time after

t_n at which $(\cdot, \hat{X}^n)$ exists B. By applying Itô's Lemma, using Lemma 2.1 and the fact that $F\varphi_n \geq 0$ on B, we obtain

$$V^n_{\theta_n} \geq (\varphi_n + \mathfrak{I})(\theta_n, \hat{X}^n_{\theta_n}, Y^n_{\theta_n}) + v_n - \varphi_n(t_n, x_n).$$

Let $2\varepsilon := \min\{|t - t_o|^2 + |x - x_o|^4,\ (t, x) \in \partial B\}$. For n large enough, the above implies

$$V^n_{\theta_n} \geq (w_{k_{n-1}} + \mathfrak{I})(\theta_n, \hat{X}^n_{\theta_n}, Y^n_{\theta_n}) + \varepsilon + \iota_n,$$

where $\iota_n := (\varphi_n - w_{k_{n-1}})(t_{n-1}, x_{n-1}) + v_n - \varphi_n(t_n, x_n)$ converges to 0. Hence, we can find n such that

$$V^n_{\theta_n} > (w_{k_{n-1}} + \mathfrak{I})(\theta_n, \hat{X}^n_{\theta_n}, Y^n_{\theta_n}).$$

Now observe that we can change the subsequence $(k_n)_{n\geq 1}$ in such a way that $k_n \geq 2k_{n-1} + 2$. Then, $v_n = w_{k_n}(t_n, x_n) - n^{-1} < w_{2k_{n-1}+2}(t_n, x_n)$, which leads to a contradiction to Proposition 2.2(ii).

b. It remains to consider the case $t_o = T$. As in Step 1., we only explain how to modify the argument used above. Let (v_n, k_n, t_n, x_n) be as in a. We now set $\varphi_n(t, x) := \varphi(t, x) + \sqrt{T - t} + |x - x_n|^4$. Since $\partial_t \varphi_n(t, x) \to -\infty$ as $t \to T$, we can find n large enough so that $F\varphi_n \geq 0$ on $[t_n, T) \times B(x_o)$ in which $B(x_o)$ is an open ball around x_o. Assume that $\varphi(T, x_o) > G(x_o) + \eta$ for some $\eta > 0$. Then, after possibly changing $B(x_o)$, we can assume that $\varphi_n(T, \cdot) \geq G + \eta$ on $B(x_o)$. We now use the same construction as in a. but with θ_n defined as the minimum between T and the first time where $\hat{X}^n$ exists $B(x_o)$. We obtain

$$V^n_{\theta_n} \geq (\varphi_n + \mathfrak{I})(\theta_n, \hat{X}^n_{\theta_n}, Y^n_{\theta_n}) + v_n - \varphi_n(t_n, x_n).$$

Let $2\varepsilon := \min\{|x - x_o|^4,\ x \in \partial B(x_o)\}$. For n large enough, the above implies

$$V^n_{\theta_n} \geq w_{k_{n-1}}(\theta_n, \hat{X}^n_{\theta_n}) \mathbf{1}_{\theta_n < T} + G(\hat{X}^n_{\theta_n}) \mathbf{1}_{\theta_n = T} + \mathfrak{I}(\hat{X}^n_{\theta_n}, Y^n_{\theta_n}) + \varepsilon \wedge \eta + \iota_n,$$

where ι_n converges to 0. By (2.3) and (H3),

$$V^n_{\theta_n} > w_{k_{n-1}}(\theta_n, \hat{X}^n_{\theta_n}) + \mathfrak{I}(\hat{X}^n_{\theta_n}, Y^n_{\theta_n}),$$

for n large enough. We conclude as in a. □

2.5.3 Comparison

In all this section, we work under the additional condition

$$\inf f > 0. \tag{2.13}$$

Direct computations (use (2.7) and Proposition 1.3) show that $\hat{F}\varphi$ is of the form

$$\hat{F}\varphi = -\partial_t\varphi - B(\cdot, f\partial_x\varphi)\partial_x\varphi - \frac{1}{2}A^2(\cdot, f\partial_x\varphi)\partial_{xx}\varphi - L(\cdot, f\partial_x\varphi) \tag{2.14}$$

where A, B and $L : (t, x, p) \in [0,T]\times\mathbb{R}\times\mathbb{R} \to \mathbb{R}$ are Lipschitz continuous functions.

Let Φ be a solution of the ordinary differential equation

$$\Phi'(t) = f(\Phi(t)),\ t \in \mathbb{R}. \tag{2.15}$$

Then, Φ is a bijection on $\mathbb{R}$ (as f is Lipschitz and $1/f$ is bounded) and the following is an immediate consequence of the definition of viscosity solutions.

Lemma 2.2. *Let v be a supersolution (resp. subsolution) of* (2.8)*. Fix $\rho > 0$. Then, $\tilde{v}$ defined by*

$$\tilde{v}(t, x) = e^{\rho t} v(t, \Phi(x)),$$

is a supersolution (resp. subsolution) of

$$\begin{aligned} 0 &= \rho\varphi - \partial_t\varphi - \left[B(\Phi, e^{-\rho t}\partial_x\varphi)/f(\Phi) - \frac{1}{2}A^2(\Phi, e^{-\rho t}\partial_x\varphi)f'(\Phi)/f(\Phi)^2\right]\partial_x\varphi \\ &\quad -\frac{1}{2}A^2(\Phi, e^{-\rho t}\partial_x\varphi)\partial_{xx}\varphi/f(\Phi)^2 - e^{\rho t}L(\Phi, e^{-\rho t}\partial_x\varphi) \end{aligned} \tag{2.16}$$

with the terminal condition

$$\varphi(T, \cdot) = e^{\rho T}G(\Phi). \tag{2.17}$$

To prove that comparison holds for (2.8), it suffices to prove that it holds for (2.16)-(2.17). For the latter, this is a consequence of the following result. It is rather standard but we provide the complete proof by lack of a precise reference.

Theorem 2.2. *Let $\mathcal{O}$ be an open subset of $\mathbb{R}$, u (resp. v) be a upper-semicontinuous subsolution (resp. lower-semicontinuous supersolution) on $[0, T) \times \mathcal{O}$ of:*

$$\rho\varphi - \partial_t\varphi - \bar{B}(\cdot, e^{-\rho t}\partial_x\varphi)\partial_x\varphi - \frac{1}{2}\bar{A}^2(\cdot, e^{-\rho t}\partial_x\varphi)\partial_{xx}\varphi - e^{\rho t}\bar{L}(\cdot, e^{-\rho t}\partial_x\varphi) = 0 \tag{2.18}$$

where $\rho > 0$ is constant, $\bar{A}, \bar{B}$ and $\bar{L} : (t, x, p) \in [0, T] \times \mathcal{O} \times \mathbb{R} \to \mathbb{R}$ are Lipschitz continuous functions. Suppose that u and v are bounded and satisfy $u \le v$ on the parabolic boundary of $[0, T) \times \mathcal{O}$, then $u \le v$ on the closure of $[0, T] \times \mathcal{O}$.

Proof. Suppose to the contrary that

$$\sup_{[0,T]\times\mathcal{O}} (u-v) > 0,$$

and define, for $n > 0$,

$$\Theta_n := \sup_{(t,x,y)\in[0,T)\times\mathcal{O}^2} \left(u(t,x) - v(t,y) - \frac{n}{2}|x-y|^2 - \frac{1}{2n}|x|^2 \right).$$

Then, there exists $\iota > 0$, such that $\Theta_n \geq \iota$ for n large enough. Since u and v are bounded and $u \leq v$ on the parabolic boundary of the domain, we can find $(t_n, x_n, y_n) \in [0,T) \times \mathcal{O}^2$ which achieves the above supremum.

As usual, we apply Ishii's Lemma combined with the sub- and super-solution properties of u and v, and the Lipschitz continuity of $\bar{A}, \bar{B}$ and $\bar{L}$ to obtain, with the notation $p_n := n(x_n - y_n)$,

$$\begin{aligned}
\rho(u(t_n,x_n) - v(t_n,y_n)) \;&\leq\; [\bar{B}(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) - \bar{B}(y_n, e^{-\rho t_n}p_n)]p_n \\
&\quad + \frac{1}{n}x_n \bar{B}(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) \\
&\quad + \frac{3n}{2}[\bar{A}(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) - \bar{A}(y_n, e^{-\rho t_n}p_n)]^2 \\
&\quad + \frac{1}{2n}\bar{A}^2(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) \\
&\quad + e^{\rho t_n}\left(\bar{L}(x_n, e^{-\rho t_n}(p_n + \frac{1}{n}x_n)) - \bar{L}(y_n, e^{-\rho t_n}p_n) \right) \\
&\leq\; C\left(n(x_n - y_n)^2 + |x_n - y_n| + \frac{1}{n}x_n^2 + \frac{1}{n} \right)
\end{aligned}$$

for some constant C which does not depend on n. In view of Lemma 2.3 below, and since $\rho > 0$ and $u(t_n,x_n) - v(t_n,y_n) \geq \Theta_n \geq \iota$, the above leads to a contradiction for n large enough. □

We conclude with the proof of the technical lemma that was used in our arguments above.

Lemma 2.3. *Let Ψ be a bounded upper-semicontinuous function on $[0,T] \times \mathbb{R}^2$, and Ψ_i, $i = 1,2$, be two non-negative lower-semicontinuous functions on $\mathbb{R}$ such that $\{\Psi_1 = 0\} = \{0\}$. For $n > 0$, set*

$$\Theta_n := \sup_{(t,x,y)\in[0,T]\times\mathbb{R}^2} \left(\Psi(t,x,y) - n\Psi_1(x-y) - \frac{1}{n}\Psi_2(x) \right)$$

and assume that there exists $(\hat{t}_n, \hat{x}_n, \hat{y}_n) \in [0,T] \times \mathbb{R}^2$ such that:

$$\Theta_n = \Psi(\hat{t}_n, \hat{x}_n, \hat{y}_n) - n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n).$$

Then, after possibly passing to a subsequence,

(i) $\lim_{n\to\infty} n\Psi_1(\hat{x}_n - \hat{y}_n) = 0$ *and* $\lim_{n\to\infty} \frac{1}{n}\Psi_2(\hat{x}_n) = 0$.

(ii) $\lim_{n\to\infty} \Theta_n = \sup_{(t,x)\in[0,T]\times\mathcal{O}} \Psi(t,x,x)$.

Proof. For later use, set $\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{\infty\}$ and note that we can extend Ψ as a bounded upper-semicontinuous function on $[0,T] \times \bar{\mathbb{R}}^2$. Set $M := \sup_{(t,x)\in[0,T]\times\mathbb{R}} \Psi(t,x,x)$, and select a sequence $(t_n, x_n)_{n\geq 1}$ such that

$$\lim_{n\to\infty} \Psi(t_n, x_n, x_n) = M \quad \text{and} \quad \lim_{n\to\infty} \frac{1}{n}\Psi_2(x_n) = 0.$$

Let C be a upper-bound for Ψ. Then,

$$\begin{aligned} C - n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n) &\geq \Psi(\hat{t}_n, \hat{x}_n, \hat{y}_n) - n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n) \\ &\geq \Psi(t_n, x_n, x_n) - \frac{1}{n}\Psi_2(x_n) \\ &\geq M - \varepsilon_n \end{aligned}$$

where $\epsilon_n \to 0$. Since Ψ_1 and Ψ_2 are non-negative, letting $n \to \infty$ in the above inequality leads to

$$\lim_{n\to\infty} \Psi_1(\hat{x}_n - \hat{y}_n) = 0$$

which implies $\lim_{n\to\infty}(\hat{x}_n - \hat{y}_n) = 0$ by the assumption $\{\Psi_1 = 0\} = \{0\}$.

After possibly passing to a subsequence, we can then assume that $\lim_{n\to\infty} \hat{x}_n = \lim_{n\to\infty} \hat{y}_n = \hat{x} \in \bar{\mathbb{R}}$ and that $\lim_{n\to\infty} \hat{t}_n = \hat{t} \in [0,T]$. Since Ψ is upper semi-continuous, the above leads to

$$\begin{aligned} & M - \liminf_{n\to\infty} \left(n\Psi_1(\hat{x}_n - \hat{y}_n) + \frac{1}{n}\Psi_2(\hat{x}_n) \right) \\ \geq\ & \Psi(\hat{t}, \hat{x}, \hat{x}) - \liminf_{n\to\infty} \left(n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n) \right) \\ \geq\ & \limsup_{n\to\infty} \left(\Psi(\hat{t}_n, \hat{x}_n, \hat{y}_n) - n\Psi_1(\hat{x}_n - \hat{y}_n) - \frac{1}{n}\Psi_2(\hat{x}_n) \right) \\ \geq\ & M, \end{aligned}$$

and our claim follows. $\square$

Remark 2.5. *It follows from the above that, whenever they are bounded, e.g. if G is bounded, then $w_* \geq w^*$. Since by construction $w_* \leq w \leq w^*$, the three functions are equal to the unique bounded viscosity solution of* (2.8)*.*

2.5.4 Smoothness

We conclude here the proof of Theorem 2.1 by showing that existence of a smooth solution holds when

$$\inf f > 0,\ G \text{ is bounded and } C^2 \text{ with } G, G', G'' \text{ Hölder continuous.} \tag{2.19}$$

Note that the assumptions $\inf f > 0$ and (H1) imply that Φ^{-1} is C^2, recall (2.15). Hence, by the same arguments as in Section 2.5.3, existence of a $C^{1,2}([0,T) \times \mathbb{R}) \cap C^0([0,T] \times \mathbb{R})$ solution to (2.16)-(2.17) implies the existence of a $C^{1,2}([0,T) \times \mathbb{R}) \cap C^0([0,T] \times \mathbb{R})$ solution to (2.8). As for (2.16)-(2.17), this is a consequence of [11, Thm 14.24], under (H1) and (2.19).

It remains to show that the solution can be taken bounded, then the comparison result of Section 2.5.3 will imply that w is this solution. Again, it suffices to work with (2.16)-(2.17). Let φ be a $C^{1,2}([0,T) \times \mathbb{R}) \cap C^0([0,T] \times \mathbb{R})$ solution of (2.16)-(2.17). Let $S^{t,x}$ be defined by

$$S_s^{t,x} = x + \int_t^s \mu_S(s, S_s^{t,x})ds + \int_t^s \sigma_S(s, S_s^{t,x})dW_s,\ s \geq t,$$

where

$$\begin{aligned} \mu_S &:= B(\Phi, e^{-\rho t}\partial_x \varphi)/f(\Phi) - \frac{1}{2}A^2(\Phi, e^{-\rho t}\partial_x \varphi)f'(\Phi)/f(\Phi)^2 \\ \sigma_S &:= A(\Phi, e^{-\rho t}\partial_x \varphi)/f(\Phi). \end{aligned}$$

Note that the coefficients of the sde may only be locally Lipschitz. However, they are bounded (recall (H1) and (2.19)), which is enough to define a solution by a standard localization procedure. Since σ_S is bounded, Itô's Lemma implies that

$$\varphi(t,x)e^{-\rho t} = \mathbb{E}\left[G(\Phi(S_T^{t,x})) + \int_t^T L(\Phi(X_s^{t,x}), e^{-\rho s}\partial_x \varphi(s, X_s^{t,x}))ds\right].$$

Since G and L are bounded, by (H1) and (2.19), φ is bounded as well. $\square$

Remark 2.6. *We refer to [10] for conditions under which additional smoothness of the solution can be proven.*

A Appendix

We report here the measurability property that was used in the course of Proposition 2.2.

In the following, $\mathcal{A}_k$ is viewed as a closed subset of the Polish space $\mathbf{L}_2^\lambda$ endowed with the usual strong norm topology $\|\cdot\|_{\mathbf{L}_2^\lambda}$.

We consider an element $\nu \in \mathcal{U}_k$ as a measurable map $\omega \in \Omega \mapsto \nu(\omega) \in \mathcal{M}_k$ where $\mathcal{M}_k$ denotes the set of non-negative Borel measures on $\mathbb{R} \times [0,T]$ with total mass less than k, endowed with the topology of weak convergence. This topology is generated by the norm

$$\|m\|_{\mathcal{M}} := \sup\{\int_{\mathbb{R}\times[0,T]} \ell(\delta,s)m(d\delta,ds) : \ell \in \mathrm{Lip}_1\},$$

in which Lip_1 denotes the class of 1-Lipschitz continuous functions bounded by 1, see e.g. [4, Proposition 7.2.2 and Theorem 8.3.2]. Then, $\mathcal{U}_k$ is a closed subset of the space $\mathbf{M}_{k,2}$ of $\mathcal{M}_k$-valued random variables. $\mathbf{M}_{k,2}$ is made complete and separable by the norm

$$\|\nu\|_{\mathbf{M}_2} := \mathbb{E}\left[\|\nu\|_{\mathcal{M}}^2\right]^{\frac{1}{2}}.$$

See e.g. [8, Chap. 5]. We endow the set of controls Γ_k with the natural product topology

$$\|\gamma\|_{\mathbf{L}_2^\lambda\times\mathbf{M}_2} := \|\vartheta\|_{\mathbf{L}_2^\lambda} + \|\nu\|_{\mathbf{M}_2}, \quad \text{for } \gamma = (\vartheta,\nu).$$

As a closed subset of the Polish space $\mathbf{L}_2^\lambda \times \mathbf{M}_{k,2}$, Γ_k is a Borel space, for each $k \geq 1$. See e.g. [3, Proposition 7.12].

The following stability result is proved by using standard estimates. In the following, we use the notation $Z = (X,Y,V)$.

Proposition A.1. *For each $k \geq 1$, there exists a real constant $c_k > 0$ such that*

$$\|Z_T^{t_1,z_1,\gamma_1} - Z_T^{t_2,z_2,\gamma_2}\|_{\mathbf{L}_2} \leq c_k\left(|t_1-t_2|^{\frac{1}{2}} + |z_1-z_2| + \|\gamma_1-\gamma_2\|_{\mathbf{L}_2^\lambda\times\mathbf{M}_2}\right),$$

for all $(t_i,z_i,\gamma_i) \in \mathrm{D}\times\Gamma_k$, $i=1,2$.

A direct consequence is the continuity of $(t,z,\gamma) \in \mathrm{D}\times\Gamma_k \mapsto Z_T^{t,z,\gamma}$, which is therefore measurable

Corollary A.1. *For each $k \geq 1$, the map $(t,z,\gamma) \in \mathrm{D}\times\Gamma_k \mapsto Z_T^{t,z,\gamma} \in \mathbf{L}_2$ is Borel-measurable.*

References

[1] F. Abergel and G. Loeper. Pricing and hedging contingent claims with liquidity costs and market impact. SSRN.

[2] G. Barles. *Solutions de viscosité des équations de Hamilton-Jacobi.* Springer Verlag, 1994.

[3] D. P. Bertsekas and S. E. Shreve. *Stochastic Optimal Control. The Discrete-Time Case.* Academic Press, New York, 1978.

[4] V. I. Bogachev and M. A. S. Ruas. *Measure theory*, volume 1. Springer, 2007.

[5] B. Bouchard, R. Elie, and N. Touzi. Stochastic target problems with controlled loss. *SIAM Journal on Control and Optimization*, 48(5):3123–3150, 2009.

[6] U. Çetin, R. A. Jarrow, and P. Protter. Liquidity risk and arbitrage pricing theory. *Finance and stochastics*, 8(3):311–341, 2004.

[7] U. Çetin, H. M. Soner, and N. Touzi. Option hedging for small investors under liquidity costs. *Finance Stoch.*, 14(3):317–341, 2010.

[8] H. Crauel. *Random probability measures on Polish spaces*, volume 11. CRC Press, 2003.

[9] R. Frey. Perfect option hedging for a large trader. *Finance and Stochastics*, 2:115–141, 1998.

[10] O. A. Ladyzhenskaia, V. Solonnikov, and N. N. Ural'tseva. *Linear and quasi-linear equations of parabolic type*, volume 23. American Mathematical Soc., 1988.

[11] G. M. Lieberman. *Second order parabolic differential equations.* World scientific, 1996.

[12] H. Liu and J. M. Yong. Option pricing with an illiquid underlying asset market. *Journal of Economic Dynamics and Control*, 29:2125–2156, 2005.

[13] G. Loeper. Option pricing with market impact and non-linear black and scholes pde's. ArxivN.

[14] H. M. Soner P. Cheridito and N. Touzi. The multi-dimensional super-replication problem under gamma constraints. *Annales de l'Institut Henri Poincaré, Série C: Analyse Non-Linéaire*, 22:633–666, 2005.

[15] P. J. Schönbucher and P. Wilmott. The feedback effects of hedging in illiquid markets. *SIAM Journal on Applied Mathematics*, 61:232–272, 2000.

[16] K. R. Sircar and G. Papanicolaou. Generalized black-scholes models accounting for increased market volatility from hedging strategies. *Applied Mathematical Finance*, 5(1):45–82, 1998.

[17] H. M. Soner and N. Touzi. Superreplication under gamma constraints. *SIAM J. Control Optim.*, 39:73–96, 2000.

[18] H. M. Soner and N. Touzi. Dynamic programming for stochastic target problems and geometric flows. *Journal of the European Mathematical Society*, 4(3):201–236, 2002.

[19] H. M. Soner and N. Touzi. Hedging under gamma constraints by optimal stopping and face-lifting. *Mathematical finance*, 17:59–80, 2007.

[20] H. M. Soner and N. Touzi. The dynamic programming equation for second order stochastic target problems. *SIAM Journal on Control and Optimization*, 48(4):2344–2365, 2009.